



- 107 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Hilton, Douglas J.  
Alexander, Warren S.  
Viney, Elizabeth M.  
Willson, Tracy A.  
Richardson, Rachael T.  
Starr, Robyn  
Nicholson, Sandra E.  
Metcalf, Donald  
Nicola, Nicos A.
- (ii) TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
- (iii) NUMBER OF SEQUENCES: 50
- (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: Scully, Scott, Murphy & Presser  
(B) STREET: 400 Garden City Plaza  
(C) CITY: Garden City, New York  
(D) STATE: New York  
(E) COUNTRY: U.S.A.  
(F) ZIP: 11530
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
(A) APPLICATION NUMBER: 08/962,560  
(B) FILING DATE: 31-OCT-1997
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: PCT/AU97/00729  
(B) FILING DATE: 31-OCT-1997
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: PO5117  
(B) FILING DATE: 14-FEB-1997
- (vii) PRIOR APPLICATION DATA:  
(A) APPLICATION NUMBER: PO 3384  
(B) FILING DATE: 01-NOV-1996
- (viii) ATTORNEY/AGENT INFORMATION:  
(A) NAME: DiGiglio, Frank S.  
(B) REGISTRATION NO: 31,346  
(C) REFERENCE/DOCKET NUMBER: 10976

- B1*  
*cont.*
- (ix) TELECOMMUNICATION INFORMATION:
    - (A) TELEPHONE: (516) 742-4343
    - (B) TELEFAX: (516) 742-4366
    - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGCCGCCC ACGTGAAGGC

20

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTCGCCAATG ACAAGACGCT

20

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1235 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 161..799

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGAGGCTCAA GCTCCGGGCG GATTCTGCGT GCCGCTCTCG CTCCTTGGGG TCTGTTGGCC

60

GGCCTGTGCC ACCCGGACGC CCGGCTCACT GCCTCTGTCT CCCCATCAG CGCAGCCCCG

120

B1

GACGCTATGG	CCCACCCCTC	CAGCTGGCCC	CTCGAGTAGG	160
ATG GTA GCA CGC AAC CAG GTG GCA GCC GAC AAT GCG ATC TCC CCG GCA	208			
Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala				
1 5 10 15				
GCA GAG CCC CGA CGG CGG TCA GAG CCC TCC TCG TCC TCG TCT TCG TCC	256			
Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser Ser				
20 25 30				
TCG CCA GCG GCC CCC GTG CGT CCC CGG CCC TGC CCG GCG GTC CCA GCC	304			
Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala				
35 40 45				
CCA GCC CCT GGC GAC ACT CAC TTC CGC ACC TTC CGC TCC CAC TCC GAT	352			
Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp				
50 55 60				
TAC CGG CGC ATC ACG CGG ACC AGC GCG CTC CTG GAC GCC TGC GGC TTC	400			
Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe				
65 70 75 80				
TAT TGG GGA CCC CTG AGC GTG CAC GGG GCG CAC GAG CGG CTG CGT GCC	448			
Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala				
85 90 95				
GAG CCC GTG GGC ACC TTC TTG GTG CGC GAC AGT CGT CAA CGG AAC TGC	496			
Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys				
100 105 110				
TTC TTC GCG CTC AGC GTG AAG ATG GCT TCG GGC CCC ACG AGC ATC CGC	544			
Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg				
115 120 125				
GTG CAC TTC CAG GCC GGC CGC TTC CAC TTG GAC GGC AGC CGC GAG ACC	592			
Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr				
130 135 140				
TTC GAC TGC CTT TTC GAG CTG CTG GAG CAC TAC GTG GCG GCG CCG CGC	640			
Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg				
145 150 155 160				
CGC ATG TTG GGG GCC CCG CTG CGC CAG CGC CGC GTG CGG CCG CTG CAG	688			
Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln				
165 170 175				
GAG CTG TGT CCG CAG CGC ATC GTG GCC GCC GTG GGT CGC GAG AAC CTG	736			
Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu				
180 185 190				

GCG CGC ATC CCT CTT AAC CCG GTA CTC CGT GAC TAC CTG AGT TCC TTC 784  
 Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
 195 200 205  
  
 CCC TTC CAG ATC TGA CCGGCTG CCGCTGTGCC GCAGCATTAA GTGGGGGCGC 836  
 Pro Phe Gln Ile  
 210  
  
 CTTATTATTT CTTATTATTA ATTATTATTA TTTTCTGGA ACCACGTGGG AGCCCTCCCC 896  
 GCCTGGGTCG GAGGGAGTGG TTGTGGAGGG TGAGATGCCT CCCACTTCTG GCTGGAGACC 956  
 TCATCCCACC TCTCAGGGGT GGGGGTGCTC CCCTCCTGGT GCTCCCTCCG GGTCCCCCCT 1016  
 GGTGTAGCA GCTTGTGTCT GGGGCCAGGA CCTGAATTCC ACTCCTACCT CTCCATGTTT 1076  
 ACATATTCCC AGTATCTTTG CACAAACCAG GGGTCGGGGA GGGTCTCTGG CTTCAATTTT 1136  
 CTGCTGTGCA GAATATCCTA TTTTATATTT TTACAGCCAG TTTAGGTAAT AAACCTTTATT 1196  
 ATGAAAGTTT TTTTAAATAA GAAAAAAAAA AAAAAAAAAA 1235

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 212 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Ala Arg Asn Gln Val Ala Ala Asp Asn Ala Ile Ser Pro Ala  
 1 5 10 15  
 Ala Glu Pro Arg Arg Arg Ser Glu Pro Ser Ser Ser Ser Ser Ser  
 20 25 30  
 Ser Pro Ala Ala Pro Val Arg Pro Arg Pro Cys Pro Ala Val Pro Ala  
 35 40 45  
 Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp  
 50 55 60  
 Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe  
 65 70 75 80  
 Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala  
 85 90 95

Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys  
 100 105 110  
 Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg  
 115 120 125  
 Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Thr  
 130 135 140  
 Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg  
 145 150 155 160  
 Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln  
 165 170 175  
 Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu  
 180 185 190  
 Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
 195 200 205  
 Pro Phe Gln Ile  
 210

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1121 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 223..819

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGATCTGTG GGTGACAGTG TCTGCGAGAG ACTTTGCCAC ACCATTCTGC CGGAATTTGG	60
AGAAAAAGAA CCAGCCGCTT CCAGTCCCCT CCCCCTCCGC CACCATTTCG GACACCCTGC	120
ACACTCTCGT TTTGGGGTAC CCTGTGACTT CCAGGCAGCA CGCGAGGTCC ACTGGCCCCA	180
GCTCGGGCGA CCAGCTGTCT GGGACGTGTT GACTCATCTC CC ATG ACC CTG CGG	234
Met Thr Leu Arg	

[illegible]

GGATAACTGC GCAGAATTCT CTCTTAAGGA CAGTTGGGCT CAGTCTAACT TAAAGGTGTG 986  
AAGATGTAGC TAGGTATTTT AAAGTTCCCC TTAGGTAGTT TTAGCTGAAT GATGCTTTCT 1046  
TTCCTATGGC TGCTCAAGAT CAAATGGCCC TTTTAAATGA AACAAAACAA AACAAAACAA 1106  
AAAAAAAAAA AAAAA 1121

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Thr Leu Arg Cys Leu Glu Pro Ser Gly Asn Gly Ala Asp Arg Thr  
1 5 10 15  
Arg Ser Gln Trp Gly Thr Ala Gly Leu Pro Glu Glu Gln Ser Pro Glu  
20 25 30  
Ala Ala Arg Leu Ala Lys Ala Leu Arg Glu Leu Ser Gln Thr Gly Trp  
35 40 45  
Tyr Trp Gly Ser Met Thr Val Asn Glu Ala Lys Glu Lys Leu Lys Glu  
50 55 60  
Ala Pro Glu Gly Thr Phe Leu Ile Arg Asp Ser Ser His Ser Asp Tyr  
65 70 75 80  
Leu Leu Thr Ile Ser Val Lys Thr Ser Ala Gly Pro Thr Asn Leu Arg  
85 90 95  
Ile Glu Tyr Gln Asp Gly Lys Phe Arg Leu Asp Ser Ile Ile Cys Val  
100 105 110  
Lys Ser Lys Leu Lys Gln Phe Asp Ser Val Val His Leu Ile Asp Tyr  
115 120 125  
Tyr Val Gln Met Cys Lys Asp Lys Arg Thr Gly Pro Glu Ala Pro Arg  
130 135 140  
Asn Gly Thr Val His Leu Tyr Leu Thr Lys Pro Leu Tyr Thr Ser Ala  
145 150 155 160  
Pro Thr Leu Gln His Phe Cys Arg Leu Ala Ile Asn Lys Cys Thr Gly  
165 170 175

B1

- 114 -

Thr Ile Trp Gly Leu Pro Leu Pro Thr Arg Leu Lys Asp Tyr Leu Glu  
180 185 190

Glu Tyr Lys Phe Gln Val  
195

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2187 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..695

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCTGGCTCC GTGCGCC ATG GTC ACC CAC AGC AAG TTT CCC GCC GCC GGG	50
Met Val Thr His Ser Lys Phe Pro Ala Ala Gly	
1 5 10	
ATG AGC CGC CCC CTG GAC ACC AGC CTG CGC CTC AAG ACC TTC AGC TCC	98
Met Ser Arg Pro Leu Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser	
15 20 25	
AAA AGC GAG TAC CAG CTG GTG GTG AAC GCC GTG CGC AAG CTG CAG GAG	146
Lys Ser Glu Tyr Gln Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu	
30 35 40	
AGC GGA TTC TAC TGG AGC GCC GTG ACC GGC GGC GAG GCG AAC CTG CTG	194
Ser Gly Phe Tyr Trp Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu	
45 50 55	
CTC AGC GCC GAG CCC GCG GGC ACC TTT CTT ATC CGC GAC AGC TCG GAC	242
Leu Ser Ala Glu Pro Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp	
60 65 70 75	
CAG CGC CAC TTC TTG ACG TTG AGC GTC AAG ACC CAG TCG GGG ACC AAG	290
Gln Arg His Phe Phe Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys	
80 85 90	
AAC CTA CGC ATC CAG TGT GAG GGG GGC AGC TTT TCG CTG CAG AGT GAC	338
Asn Leu Arg Ile Gln Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp	
95 100 105	



B1

- 115 -

CCC CGA AGC ACG CAG CCA GTT CCC CGC TTC GAC TGT GTA CTC AAG CTG Pro Arg Ser Thr Gln Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu 110 115 120	386
GTG CAC CAC TAC ATG CCG CCT CCA GGG ACC CCC TCC TTT TCT TTG CCA Val His His Tyr Met Pro Pro Pro Gly Thr Pro Ser Phe Ser Leu Pro 125 130 135	434
CCC ACG GAA CCC TCG TCC GAA GTT CCG GAG CAG CCA CCT GCC CAG GCA Pro Thr Glu Pro Ser Ser Glu Val Pro Glu Gln Pro Pro Ala Gln Ala 140 145 150 155	482
CTC CCC GGG AGT ACC CCC AAG AGA GCT TAC TAC ATC TAT TCT GGG GGC Leu Pro Gly Ser Thr Pro Lys Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly 160 165 170	530
GAG AAG ATT CCG CTG GTA CTG AGC CGA CCT CTC TCC TCC AAC GTG GCC Glu Lys Ile Pro Leu Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala 175 180 185	578
ACC CTC CAG CAT CTT TGT CGG AAG ACT GTC AAC GGC CAC CTG GAC TCC Thr Leu Gln His Leu Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser 190 195 200	626
TAT GAG AAA GTG ACC CAG CTG CCT GGA CCC ATT CGG GAG TTC CTG GAT Tyr Glu Lys Val Thr Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp 205 210 215	674
CAG TAT GAT GCT CCA CTT TAAGGAGCAA AAGGGTCAGA GGGGGGCCTG Gln Tyr Asp Ala Pro Leu 220 225	722
GGTCGGTCGG TCGCCTCTCC TCCGAGGCAC ATGGCACAAG CACAAAAATC CAGCCCCAAC	782
GGTCGGTAGC TCCCAGTGAG CCAGGGGCAG ATTGGCTTCT TCCTCAGGCC CTCCACTCCC	842
GCAGAGTAGA GCTGGCAGGA CCTGGAATTC GTCTGAGGGG AGGGGGAGCT GCCACCTGCT	902
TTCCCCCCTC CCCCAGCTCC AGCTTCTTTC AAGTGGAGCC AGCCGGCCTG GCCTGGTG	962
ACAATACCTT TGACAAGCGG ACTCTCCCCT CCCCTTCCTC CACACCCCCT CTGCTTCCCA	1022
AGGGAGGTGG GGACACCTCC AAGTGTGAA CTTAGAACTG CAAGGGGAAT CTTCAAACCTT	1082
TCCCCTGGA ACTTGTTTGC GCTTTGATTT GGTGTTGATCA AGAGCAGGCA CCTGGGGGAA	1142
GGATGGAAGA GAAAAGGGTG TGTGAAGGGT TTTTATGCTG GCCAAAGAAA TAACCACTCC	1202
CACTGCCCAA CCTAGGTGAG GAGTGGTGGC TCCTGGCTCT GGGGAGAGTG GCAAGGGGTG	1262
ACCTGAAGAG AGCTATACTG GTGCCAGGCT CCTCTCCATG GGGCAGCTAA TGAAACCTCG	1322

*Dist*

- 116 -

CAGATCCCTT GCACCCAGAC ACCCTCCCCG TTGTGAAGAG GCAGTAGCAT TTAGAAGGGA	1382
GACAGATGAG GCTGGTGAGC TGGCCGCCTT TTCCAACACC GAAGGGAGGC AGATCAACAG	1442
ATGAGCCATC TTGGAGCCCA GGTTCCTCCCT GGAGCAGATG GAGGGTTCTG CTTTGTCTCT	1502
CCTATGTGGG GCTAGGAGAC TCGCCTTAAA TGCCCTCTGT CCCAGGGATG GGGATTGGCA	1562
CACAAGGAGC CAAACACAGC CAATAGGCAG AGAGTTGAGG GATTCACCCA GGTGGCTACA	1622
GGCCAGGGGA AGTGGCTGCA GGGGAGAGAC CCAGTCACTC CAGGAGACTC CTGAGTTAAC	1682
ACTGGGAAGA CATTGGCCAG TCCTAGTCAT CTCTCGGTCA GTAGGTCCGA GAGCTTCCAG	1742
GCCCTGCACA GCCCTCCTTT CTCACCTGGG GGGAGGCAGG AGGTGATGGA GAAGCCTTCC	1802
CATGCCGCTC ACAGGGGCCT CACGGGAATG CAGCAGCCAT GCAATTACCT GGAAGTGGTC	1862
CTGTGTTGGG GAGAAACAAG TTTTCTGAAG TCAGGTATGG GGCTGGGTGG GGCAGCTGTG	1922
TGTTGGGGTG GCTTTTTTCT CTCTGTTTTG AATAATGTTT ACAATTTGCC TCAATCACTT	1982
TTATAAAAT CCACCTCCAG CCCGCCCTC TCCCCACTCA GGCCTTCGAG GCTGTCTGAA	2042
GATGCTTGAA AACTCAACC AAATCCCAGT TCAACTCAGA CTTTGCACAT ATATTTATAT	2102
TTATACTCAG AAAAGAAACA TTTCAATAAT TTATAATAAA AGAGCACTAT TTTTAAATGA	2162
AAAAAAAAA AAAAAAAAAA AAAAA	2187

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 225 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Thr His Ser Lys Phe Pro Ala Ala Gly Met Ser Arg Pro Leu	
1 5 10 15	
Asp Thr Ser Leu Arg Leu Lys Thr Phe Ser Ser Lys Ser Glu Tyr Gln	
20 25 30	
Leu Val Val Asn Ala Val Arg Lys Leu Gln Glu Ser Gly Phe Tyr Trp	
35 40 45	
Ser Ala Val Thr Gly Gly Glu Ala Asn Leu Leu Ser Ala Glu Pro	
50 55 60	

B1

- 117 -

Ala Gly Thr Phe Leu Ile Arg Asp Ser Ser Asp Gln Arg His Phe Phe  
65 70 75 80

Thr Leu Ser Val Lys Thr Gln Ser Gly Thr Lys Asn Leu Arg Ile Gln  
85 90 95

Cys Glu Gly Gly Ser Phe Ser Leu Gln Ser Asp Pro Arg Ser Thr Gln  
100 105 110

Pro Val Pro Arg Phe Asp Cys Val Leu Lys Leu Val His His Tyr Met  
115 120 125

Pro Pro Pro Gly Thr Pro Ser Phe Ser Leu Pro Pro Thr Glu Pro Ser  
130 135 140

Ser Glu Val Pro Glu Gln Pro Pro Ala Gln Ala Leu Pro Gly Ser Thr  
145 150 155 160

Pro Lys Arg Ala Tyr Tyr Ile Tyr Ser Gly Gly Glu Lys Ile Pro Leu  
165 170 175

Val Leu Ser Arg Pro Leu Ser Ser Asn Val Ala Thr Leu Gln His Leu  
180 185 190

Cys Arg Lys Thr Val Asn Gly His Leu Asp Ser Tyr Glu Lys Val Thr  
195 200 205

Gln Leu Pro Gly Pro Ile Arg Glu Phe Leu Asp Gln Tyr Asp Ala Pro  
210 215 220

Leu  
225

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1094 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCCGGCTGG CCCCTTCTGT AGGATGGTAG CACACAACCA GGTGGCAGCC GACAATGCAG 60

TCTCCACAGC AGCAGAGCCC CGACGGCGGC CAGAACCTTC CTCCTCTTCC TCCTCCTCGC 120

CCGCGGCCCC CGCGCGCCCG CGGCCGTGCC CCGCGGTCCC GGCCCCGGCC CCCGGCGACA 180

B1

- 118 -

```

CGCACTTCCG CACATTCCGT TCGCACGCCG ATTACCGGCG CATCACGCGC GCCAGCGCGC      240
TCCTGGACGC CTGCGGATTC TACTGGGGGC CCCTGAGCGT GCACGGGGCG CACGAGCGGC      300
TGCGCGCCGA GCCCGTGGGC ACCTTCCTGG TGC GCGACAG CCGCCAGCGG AACTGCTTTT      360
TCGCCCTTAG CGTGAAGATG GCCTCGGGAC CCACGAGCAT CCGCGTGCAC TTTCAGGCCG      420
GCCGCTTTCA CCTGGATGGC AGCCGCGAGA GCTTCGACTG CCTCTTCGAG CTGCTGGAGC      480
ACTACGTGGC GGC GCCGCGC CGCATGCTGG GGGCCCCGCT GCGCCAGCGC CGCGTGCGGC      540
CGCTGCAGGA GCTGTGCCGC CAGCGCATCG TGGCCACCGT GGGCCGCGAG AACCTGGCTC      600
GCATCCCCCT CAACCCCGTC CTCCGCGACT ACCTGAGCTC CTTCCCCTTC CAGATTTGAC      660
CGGCAGCGCC CGCCGTGCAC GCAGCATTAA CTGGGATGCC GTGTTATTTT GTTATTACTT      720
GCCTGGAACC ATGTGGGTAC CCTCCCCGGC CTGGGTTGGA GGGAGCGGAT GGGTGTAGGG      780
GCGAGGCGCC TCCCGCCCTC GGCTGGAGAC GAGGCCGCAG ACCCCTTCTC ACCTCTTGAG      840
GGGGTCCTCC CCCTCCTGGT GCTCCCTCTG GGTCCCCCTG GTTGTGTAG CAGCTTAACT      900
GTATCTGGAG CCAGGACCTG AACTCGCACC TCCTACCTCT TCATGTTTAC ATATACCCAG      960
TATCTTTGCA CAAACCAGGG GTTGGGGGAG GGTCTCTGGC TTTATTTTTC TGCTGTGCAG     1020
AATCCTATTT TATATTTTTT AAAGTCAGTT TAGGTAATAA ACTTTATTAT GAAAGTTTTT     1080
TTTTTTAAAA AAAA                                                                1094

```

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Val Ala His Asn Gln Val Ala Ala Asp Asn Ala Val Ser Thr Ala
1           5           10           15
Ala Glu Pro Arg Arg Arg Pro Glu Pro Ser Ser Ser Ser Ser Ser Ser
20           25           30
Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Ala Val Pro Ala Pro
35           40           45

```

Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ala Asp Tyr  
50 55 60

Arg Arg Ile Thr Arg Ala Ser Ala Leu Leu Asp Ala Cys Gly Phe Tyr  
65 70 75 80

Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ala Glu  
85 90 95

Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys Phe  
100 105 110

Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg Val  
115 120 125

His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Ser Arg Glu Ser Phe  
130 135 140

Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg Arg  
145 150 155 160

Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln Glu  
165 170 175

Leu Cys Arg Gln Arg Ile Val Ala Thr Val Gly Arg Glu Asn Leu Ala  
180 185 190

Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe Pro  
195 200 205

Phe Gln Ile  
210

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2807 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGAAACCGAG GCGGGGAGAC CAGGAGGCCT TGGCCTCAGA GCTTCAGAGT CGCGTGGCAG	60
CAAACAGAGA AACCTGTAGA GGGCAGTGTG CGTCACTTAG CTCAGGGAAG CTGCACGCGA	120
AACTACCCG CCTTCATTCA TAAACATCGT CAGCTAGGCA CCTACTCCTG GGCTTTCAGG	180
ACAAACTGAA TCACGAAACC ACAGTGTCTT TAAAATAGGT CTGACCGCCT GAATCCCTGG	240

B1

- 120 -

CCAAGGTGTG	TACGGGGCAT	GGGAGCCCTT	GTGCAGAGAT	GCTTGCAGGA	GCCTTGAGGG	300
GCTCTGTAAG	ACAGAGGCTA	GGAAGACAAA	GTTGGGGGCT	ACAGCTTCTT	GTCCTGCCCCG	360
GGGCCTCAGT	TTCTTCGGTT	GCCCACGTAG	GAGTGCAGAG	AGTCCAGCCC	CTGGGGACCC	420
AACCCAACCC	CGCCCAGTTT	CCGAGGAACT	CGTCCGGGAG	CGGGGGCGCC	CCTCCCGCAC	480
CGCCTTAGGC	TTCTTTTGAA	GCCTCTGCGG	TCAGGCCACC	GCTTCCTGGG	AAGCCCAAGC	540
CAAGGCCAGG	CCGAGTGGCC	AACGGGAGGG	GCCCCGCGCC	GATTCTGGAG	GAGGGCGGCG	600
GCCCCACAGG	TCTCCAGGGC	TGGCTAGCCG	GGCTCCTAGA	GCGGAGACTG	CCAAGGCCTT	660
CGGGTCCTGG	GCAGGAAGGA	TCCTGGCAGG	GAGGAGTTGC	TTGGGGGGTG	GGGGGGAAAG	720
GCTCCAGGCG	CGGTGGAGCT	CTGACCAGGA	GAATGCACAC	ACTCGGAGGG	GAGGAGGCGT	780
GTCAGCCCCA	AGCTAGCATC	CCACCCGGGG	AGCAGCGATG	TGGGGCGAAG	GTAGCCAGAG	840
CAAAAGAGCA	GGCACCAGGT	GACACGAAAC	AGAAGATTCC	GGGTAGAGCC	AGAACCCAG	900
AAGTCCCATT	CAGGGAAGGT	GCGAGGCGAG	AACGAGTTAG	GTGGACCCTC	TCCAGGGGCA	960
GCCAAAGAAA	TCTAAAGAGA	ACCCGAAGGA	CTTGCCGGAA	AGAGAAACCG	AAAGCGGCGG	1020
TGGGCGGGAT	CGGTGGGCGG	GGCCTCCCTG	GTTTAAGAGC	TTGATGCAGG	GGCGGGCAGC	1080
AGCAGAGAGA	ACTGCGGCCG	TGGCAGCGGC	ACGGCTCCCG	GCCCCGGAGC	ATGCGCGACA	1140
GCAGCCCCGG	AACCCCCAGC	CGCGGGGCCC	CGCGTCCCGC	CGCCAGGTGA	GCCGAGGCAG	1200
CTGCGAAGGA	GCAGGCGGGA	GGGGATGGGA	GGAAGGGGAG	CAGAGCCTGG	CAGGACTATC	1260
CTCGCAGACT	GCATGGCGGG	GTCGTGGATG	CTATGCCTCT	GGCGCCCGCC	CCACCGGCTG	1320
GCCCAGGCGG	CCCCTCGCGC	GCGCGGGGCG	CCGTCAGCCC	CTCCTCTCCG	GCCCTGAGCC	1380
CGGATCGTCC	GCCCCGGGTT	CAGTTCCCGG	CGTGGCCAGT	AGGCGGCAAC	CGCGAGGCGG	1440
CAAGCCACCC	AGCGGGGACG	GCCTGGAGTC	GGGCCCCTCT	CCACGCCCCC	TTCTCCACGC	1500
GCGCGGGGAG	GCAGGGCTCC	ACCGCCAGTC	TGGAAGGGTT	CCACATACAG	GAACGGCCTA	1560
CTTCGCAGAT	GAGCCCACCG	AGGCTCAGGC	TCCGGGCGGA	TTCTGCGTGT	CACCCTCGCT	1620
CCTTGGGGTC	CGCTGGCCGG	CCTGTGCCAC	CCGGACGCC	GGTTCACTGC	CTCTGTCTCC	1680
CCCATCAGCG	CAGCCCCGGA	CGCTATGGCC	CACCCCTCCA	GCTGGCCCCT	CGAGTAGGAT	1740
GGTAGCACGT	AACCAGGTGG	AAGCCGACAA	TGCGATCTCC	CCGGCATCAG	AGCCCCGACG	1800

B1

- 121 -

GCGGCCAGAG	CCATCCTCGT	CCTCGTCTTC	GTCTCGCCG	GCGGCCCCGG	CGCGTCCCCG	1860
GCCCTGCCCC	GTGGTCCCCG	CCCCGGCTCC	GGGCGACACT	CACTTCCGCA	CCTTCCGCTC	1920
CCACTCTGAT	TACCGGCGCA	TCACGCGGAC	CAGCGCTCTC	CTGGACGCCT	GCGGCTTCTA	1980
CTGGGGACCC	CTGAGCGTGC	ATGGGGCGCA	CGAACGGCTG	CGTTCCGAAC	CCGTGGGCAC	2040
CTTCTTGGTG	CGCGACAGTC	GCCAGCGGAA	CTGCTTCTTC	GCGCTCAGCG	TGAAGATGGC	2100
TTCGGGCCCC	ACGAGCATTC	GTGTGCACTT	CCAGGCCGGC	CGCTTCCACC	TGGACGGCAA	2160
CCGCGAGACC	TTCGACTGCC	TCTTCGAGCT	GCTGGAGCAC	TACGTGGCGG	CGCCGCGCCG	2220
CATGTTGGGG	GCCCCACTGC	GCCAGCGCCG	CGTGGGGCCG	CTGCAGGAGC	TGTGTCGCCA	2280
GCGCATCGTG	GCCGCCGTGG	GTCGCGAGAA	CCTGGCACGC	ATCCCTCTTA	ACCCGGTACT	2340
CCGTGACTAC	CTGAGTTCCT	TCCCCTTCCA	GATCTGACCG	GCTGCCGCCG	TGCCCCGAGA	2400
ATTAAGTGGG	AGCGCCTTAT	TATTTCTTAT	TATTAATTAT	TATTATTTT	CTGGAACCAC	2460
GTGGGAGCCC	TCCCCGCCTA	GGTCGGAGGG	AGTGGGTGTG	GAGGGTGAGA	TCCCTCCCAC	2520
TTCTGGCTGG	AGACCTTATC	CCGCCTCTCG	GGGGGCCTCC	CCTCCTGGTG	CTCCCTCCCG	2580
GTCCCCCTGG	TTGTAGCAGC	TTGTGTCTGG	GGCCAGGACC	TGAACTCCAC	GCCTACCTCT	2640
CCATGTTTAC	ATGTTCCCAG	TATCTTTGCA	CAAACCAGGG	GTGGGGGAGG	GTCTCTGGCT	2700
TCATTTTTCT	GCTGTGCAGA	ATATTCTATT	TTATATTTTT	ACATCCAGTT	TAGATAATAA	2760
ACTTTATTAT	GAAAGTTTTT	TTTTTTAAAG	AAACAAAGAT	TTCTAGA		2807

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 212 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met	Val	Ala	Arg	Asn	Gln	Val	Glu	Ala	Asp	Asn	Ala	Ile	Ser	Pro	Ala
1				5					10					15	
Ser	Glu	Pro	Arg	Arg	Arg	Pro	Glu	Pro	Ser	Ser	Ser	Ser	Ser	Ser	Ser
			20					25						30	

Ser Pro Ala Ala Pro Ala Arg Pro Arg Pro Cys Pro Val Val Pro Ala  
 35 40 45  
 Pro Ala Pro Gly Asp Thr His Phe Arg Thr Phe Arg Ser His Ser Asp  
 50 55 60  
 Tyr Arg Arg Ile Thr Arg Thr Ser Ala Leu Leu Asp Ala Cys Gly Phe  
 65 70 75 80  
 Tyr Trp Gly Pro Leu Ser Val His Gly Ala His Glu Arg Leu Arg Ser  
 85 90 95  
 Glu Pro Val Gly Thr Phe Leu Val Arg Asp Ser Arg Gln Arg Asn Cys  
 100 105 110  
 Phe Phe Ala Leu Ser Val Lys Met Ala Ser Gly Pro Thr Ser Ile Arg  
 115 120 125  
 Val His Phe Gln Ala Gly Arg Phe His Leu Asp Gly Asn Arg Glu Thr  
 130 135 140  
 Phe Asp Cys Leu Phe Glu Leu Leu Glu His Tyr Val Ala Ala Pro Arg  
 145 150 155 160  
 Arg Met Leu Gly Ala Pro Leu Arg Gln Arg Arg Val Arg Pro Leu Gln  
 165 170 175  
 Glu Leu Cys Arg Gln Arg Ile Val Ala Ala Val Gly Arg Glu Asn Leu  
 180 185 190  
 Ala Arg Ile Pro Leu Asn Pro Val Leu Arg Asp Tyr Leu Ser Ser Phe  
 195 200 205  
 Pro Phe Gln Ile  
 210

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1611 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 263..1529

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:



CGAATTCCGG	GCGGGCTGTG	TGAGTCTGTG	AGTGGAAGGC	GCGCCGGCTC	TTTTGTCTGA	60
GTGTGACCCG	GTGGCTTTGT	TCCAGGCATT	CCGGTGATTT	CCTCCGGGCA	GTCCGCAGAA	120
GCCGCAGCGG	CCGCCCCGCG	TCTCTCTGCA	GTCTCCACAC	CCGGGAGAGC	CTGAGCCCCG	180
GTCACGCCCC	TCAGCCCCCG	CTGAGTCCCT	TCTCTGTTGT	CGCGTCCGAA	TCGAGTCCCC	240
GGAATCAGAC	GGTGCCCCAT	AG	ATG GCC AGC TTT CCC CCG AGG GTT AAC GAG			292
			Met Ala Ser Phe Pro Pro Arg Val Asn Glu			
			1 5 10			
AAA GAG ATC GTG AGA TCA CGT ACT ATA GGG GAA CTC TTG GCT CCA GCA						340
Lys Glu Ile Val Arg Ser Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala						
	15		20		25	
GCT CCT TTT GAC AAG AAA TGT GGT GGT GAG AAC TGG ACG GTT GCT TTT						388
Ala Pro Phe Asp Lys Lys Cys Gly Gly Glu Asn Trp Thr Val Ala Phe						
	30		35		40	
GCT CCT GAT GGT TCC TAC TTT GCG TGG TCA CAA GGA TAT CGC ATA GTG						436
Ala Pro Asp Gly Ser Tyr Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val						
	45		50		55	
AAG CTT GTC CCG TGG TCC CAG TGC CGT AAG AAC TTT CTT TTG CAT GGT						484
Lys Leu Val Pro Trp Ser Gln Cys Arg Lys Asn Phe Leu Leu His Gly						
	60		65		70	
TCC AAA AAT GTT ACC AAT TCA AGC TGT CTA AAA TTG GCA AGA CAA AAC						532
Ser Lys Asn Val Thr Asn Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn						
	75		80		85	90
AGT AAT GGT GGT CAG AAA AAC AAG CCT CCT GAG CAC GTT ATA GAC TGT						580
Ser Asn Gly Gly Gln Lys Asn Lys Pro Pro Glu His Val Ile Asp Cys						
		95		100		105
GGA GAC ATA GTC TGG AGT CTT GCT TTT GGG TCT TCA GTT CCA GAA AAA						628
Gly Asp Ile Val Trp Ser Leu Ala Phe Gly Ser Ser Val Pro Glu Lys						
	110		115		120	
CAG AGT CGT TGC GTT AAT ATA GAA TGG CAT CGG TTC CGA TTT GGA CAG						676
Gln Ser Arg Cys Val Asn Ile Glu Trp His Arg Phe Arg Phe Gly Gln						
	125		130		135	
GAT CAG CTA CTC CTT GCC ACA GGA TTA AAC AAT GGT CGC ATC AAA ATC						724
Asp Gln Leu Leu Leu Ala Thr Gly Leu Asn Asn Gly Arg Ile Lys Ile						
	140		145		150	

B1

- 124 -

TGG	GAT	GTA	TAT	ACA	GGA	AAA	CTC	CTC	CTT	AAT	TTG	GTA	GAC	CAC	ATT	772
Trp	Asp	Val	Tyr	Thr	Gly	Lys	Leu	Leu	Leu	Asn	Leu	Val	Asp	His	Ile	
155					160					165					170	
GAA	ATG	GTT	AGA	GAT	TTA	ACT	TTT	GCT	CCA	GAT	GGG	AGC	TTA	CTC	CTT	820
Glu	Met	Val	Arg	Asp	Leu	Thr	Phe	Ala	Pro	Asp	Gly	Ser	Leu	Leu	Leu	
				175					180					185		
GTA	TCA	GCT	TCA	AGA	GAC	AAA	ACT	CTA	AGA	GTG	TGG	GAC	CTG	AAA	GAT	868
Val	Ser	Ala	Ser	Arg	Asp	Lys	Thr	Leu	Arg	Val	Trp	Asp	Leu	Lys	Asp	
			190					195					200			
GAT	GGA	AAC	ATG	GTG	AAA	GTA	TTG	CGG	GCA	CAT	CAG	AAT	TGG	GTG	TAC	916
Asp	Gly	Asn	Met	Val	Lys	Val	Leu	Arg	Ala	His	Gln	Asn	Trp	Val	Tyr	
		205					210					215				
AGT	TGT	GCA	TTC	TCT	CCC	GAC	TGT	TCT	ATG	CTG	TGT	TCA	GTG	GGC	GCC	964
Ser	Cys	Ala	Phe	Ser	Pro	Asp	Cys	Ser	Met	Leu	Cys	Ser	Val	Gly	Ala	
	220					225					230					
AGT	AAA	GCA	GTT	TTC	CTT	TGG	AAT	ATG	GAT	AAA	TAC	ACC	ATG	ATT	AGG	1012
Ser	Lys	Ala	Val	Phe	Leu	Trp	Asn	Met	Asp	Lys	Tyr	Thr	Met	Ile	Arg	
235					240					245					250	
AAG	CTG	GAA	GGT	CAT	CAC	CAT	GAT	GTT	GTA	GCT	TGT	GAC	TTT	TCT	CCT	1060
Lys	Leu	Glu	Gly	His	His	His	Asp	Val	Val	Ala	Cys	Asp	Phe	Ser	Pro	
				255					260					265		
GAT	GGA	GCA	TTG	CTA	GCT	ACT	GCA	TCC	TAT	GAC	ACT	CGT	GTG	TAT	GTC	1108
Asp	Gly	Ala	Leu	Leu	Ala	Thr	Ala	Ser	Tyr	Asp	Thr	Arg	Val	Tyr	Val	
			270					275					280			
TGG	GAT	CCA	CAC	AAT	GGA	GAC	CTT	CTG	ATG	GAG	TTT	GGG	CAC	CTG	TTT	1156
Trp	Asp	Pro	His	Asn	Gly	Asp	Leu	Leu	Met	Glu	Phe	Gly	His	Leu	Phe	
		285					290					295				
CCC	TCG	CCC	ACT	CCA	ATA	TTT	GCT	GGA	GGA	GCA	AAT	GAC	CGA	TGG	GTG	1204
Pro	Ser	Pro	Thr	Pro	Ile	Phe	Ala	Gly	Gly	Ala	Asn	Asp	Arg	Trp	Val	
	300					305					310					
AGA	GCT	GTG	TCT	TTC	AGT	CAT	GAT	GGA	CTG	CAT	GTT	GCC	AGC	CTT	GCT	1252
Arg	Ala	Val	Ser	Phe	Ser	His	Asp	Gly	Leu	His	Val	Ala	Ser	Leu	Ala	
315					320					325					330	
GAT	GAT	AAA	ATG	GTG	AGG	TTC	TGG	AGA	ATC	GAT	GAG	GAT	TGT	CCG	GTA	1300
Asp	Asp	Lys	Met	Val	Arg	Phe	Trp	Arg	Ile	Asp	Glu	Asp	Cys	Pro	Val	
				335					340					345		
CAA	GTT	GCA	CCT	TTG	AGC	AAT	GGT	CTT	TGC	TGT	GCC	TTT	TCT	ACT	GAT	1348
Gln	Val	Ala	Pro	Leu	Ser	Asn	Gly	Leu	Cys	Cys	Ala	Phe	Ser	Thr	Asp	
			350					355					360			

B1

- 125 -

GGC AGT GTT TTA GCT GCT GGG ACA CAT GAT GGA AGT GTG TAT TTT TGG	1396
Gly Ser Val Leu Ala Ala Gly Thr His Asp Gly Ser Val Tyr Phe Trp	
365 370 375	
GCC ACT CCA AGG CAA GTC CCT AGC CTT CAA CAT ATA TGT CGC ATG TCA	1444
Ala Thr Pro Arg Gln Val Pro Ser Leu Gln His Ile Cys Arg Met Ser	
380 385 390	
ATC CGA AGA GTG ATG TCC ACC CAA GAA GTC CAA AAA CTG CCT GTT CCT	1492
Ile Arg Arg Val Met Ser Thr Gln Glu Val Gln Lys Leu Pro Val Pro	
395 400 405 410	
TCC AAA ATA TTG GCG TTT CTC TCC TAC GCG GGT TAG A CTGAAGACTG	1539
Ser Lys Ile Leu Ala Phe Leu Ser Tyr Arg Gly	
415 420	
CCTTTCCTGG TAGGCCTGCC AGACAGAGCG OCCTTTACAA GACACACCTC AAGCTTTACC	1599
TCGTGCCGAA TT	1611

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 421 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Ser Phe Pro Pro Arg Val Asn Glu Lys Glu Ile Val Arg Ser	
1 5 10 15	
Arg Thr Ile Gly Glu Leu Leu Ala Pro Ala Ala Pro Phe Asp Lys Lys	
20 25 30	
Cys Gly Gly Glu Asn Trp Thr Val Ala Phe Ala Pro Asp Gly Ser Tyr	
35 40 45	
Phe Ala Trp Ser Gln Gly Tyr Arg Ile Val Lys Leu Val Pro Trp Ser	
50 55 60	
Gln Cys Arg Lys Asn Phe Leu Leu His Gly Ser Lys Asn Val Thr Asn	
65 70 75 80	
Ser Ser Cys Leu Lys Leu Ala Arg Gln Asn Ser Asn Gly Gly Gln Lys	
85 90 95	
Asn Lys Pro Pro Glu His Val Ile Asp Cys Gly Asp Ile Val Trp Ser	
100 105 110	

B1

- 126 -

Leu	Ala	Phe	Gly	Ser	Ser	Val	Pro	Glu	Lys	Gln	Ser	Arg	Cys	Val	Asn		
		115					120					125					
Ile	Glu	Trp	His	Arg	Phe	Arg	Phe	Gly	Gln	Asp	Gln	Leu	Leu	Leu	Ala		
	130					135					140						
Thr	Gly	Leu	Asn	Asn	Gly	Arg	Ile	Lys	Ile	Trp	Asp	Val	Tyr	Thr	Gly		
145					150					155					160		
Lys	Leu	Leu	Leu	Asn	Leu	Val	Asp	His	Ile	Glu	Met	Val	Arg	Asp	Leu		
				165					170					175			
Thr	Phe	Ala	Pro	Asp	Gly	Ser	Leu	Leu	Leu	Val	Ser	Ala	Ser	Arg	Asp		
			180					185						190			
Lys	Thr	Leu	Arg	Val	Trp	Asp	Leu	Lys	Asp	Asp	Gly	Asn	Met	Val	Lys		
		195					200					205					
Val	Leu	Arg	Ala	His	Gln	Asn	Trp	Val	Tyr	Ser	Cys	Ala	Phe	Ser	Pro		
	210					215					220						
Asp	Cys	Ser	Met	Leu	Cys	Ser	Val	Gly	Ala	Ser	Lys	Ala	Val	Phe	Leu		
225					230					235					240		
Trp	Asn	Met	Asp	Lys	Tyr	Thr	Met	Ile	Arg	Lys	Leu	Glu	Gly	His	His		
				245					250					255			
His	Asp	Val	Val	Ala	Cys	Asp	Phe	Ser	Pro	Asp	Gly	Ala	Leu	Leu	Ala		
			260					265					270				
Thr	Ala	Ser	Tyr	Asp	Thr	Arg	Val	Tyr	Val	Trp	Asp	Pro	His	Asn	Gly		
		275					280					285					
Asp	Leu	Leu	Met	Glu	Phe	Gly	His	Leu	Phe	Pro	Ser	Pro	Thr	Pro	Ile		
	290					295				300							
Phe	Ala	Gly	Gly	Ala	Asn	Asp	Arg	Trp	Val	Arg	Ala	Val	Ser	Phe	Ser		
305					310					315					320		
His	Asp	Gly	Leu	His	Val	Ala	Ser	Leu	Ala	Asp	Asp	Lys	Met	Val	Arg		
				325					330					335			
Phe	Trp	Arg	Ile	Asp	Glu	Asp	Cys	Pro	Val	Gln	Val	Ala	Pro	Leu	Ser		
			340				345						350				
Asn	Gly	Leu	Cys	Cys	Ala	Phe	Ser	Thr	Asp	Gly	Ser	Val	Leu	Ala	Ala		
		355					360					365					
Gly	Thr	His	Asp	Gly	Ser	Val	Tyr	Phe	Trp	Ala	Thr	Pro	Arg	Gln	Val		
	370					375					380						

BL

- 127 -

Pro Ser Leu Gln His Ile Cys Arg Met Ser Ile Arg Arg Val Met Ser  
385 390 395 400

Thr Gln Glu Val Gln Lys Leu Pro Val Pro Ser Lys Ile Leu Ala Phe  
405 410 415

Leu Ser Tyr Arg Gly  
420

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 783 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTGTCTTCCT CCGCAGCGCG AGGCTGGGTA CAGGGTCTAT TGTCTGTGGT TGA	60
CTTTGGTCTG AGGCCTTCGG GAGCTTTCCC GAGGCAGTTA GCAGAAGCCG CAGCGACCGC	120
CCCCGCCCGT CTCCTCTGTC CCTGGGCCCC GGAGACAAAC TTGGCGTCAC GCCCTCAGCG	180
GTCGCCACTC TCTTCTCTGT TGTTGGGTCC GCATCGTATT CCCGGAATCA GACGGTGCCC	240
CATAGATGGC CAGCTTTCCC CCGAGGGTCA ACGAGAAAGA GATCGTGAGA TCACGTACTA	300
TAGGTGAACT TTTAGCTCCT GCAGCTCCTT TTGACAAGAA ATGTGGTCGT GAAAATTGGA	360
CTGTTGCTTT TGCTCCAGAT GGTTCATACT TTGCTTGGTC ACAAGGACAT CGCACAGTAA	420
AGCTTGTTCC GTGGTCCCAG TGCCTTCAGA ACTTCTCTT GCATGGCACC AAGAATGTTA	480
CCAATTCAAG CAGTTTAAGA TTGCCAAGAC AAAATAGTGA TGGTGGTCAG AAAAATAAGC	540
CTCGTGACAT ATTATAGACT GTGGAGATAT AGTCTGGAGT CTTGCTTTTG GGTCATCAGT	600
TCCAGAAAAA CAGAGTCGCT GTGTAAATAT AGAATGGCAT CGCTTCAGAT TTGGACAAGA	660
TCAGCTACTT CTTGCTACAG GGTGAACAA TGGGCGTATC AAAATATGGG ATGTATATCA	720
GGAAACTCCT CCTTAACTTG GTAGATCATA CTGAAGTGGT CAGAGATTTA ACTTTTGCTC	780
CAG	783

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

B1

- 128 -

- (A) LENGTH: 1122 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTCTGTATGT CTGAATGAAG CTATAACATT TGCCTTTTTTA TTGCAGGTTT TCCTTTGGAA	60
TATGGATAAA TACACCATGA TACGGAAACT AGAAGGACAT CACCATGATG TGGTAGCTTG	120
TGACTTTTCT CCTGATGGAG CATTACTGGC TACTGCATCT TATGATACTC GAGTATATAT	180
CTGGGATCCA CATAATGGAG ACATTCTGAT GGAAATTGGG CACCTGTTTC CCCACCTAC	240
TCCAATATTT GCTGGAGGAG CAAATGACCG GTGGGTACGA TCTGTATCTT TTAGCCATGA	300
TGGACTGCAT GTTGCAAGCC TTGCTGATGA TAAAATGGTG AGGTTCTGGA GAATTGATGA	360
GGATTATCCA GTGCAAGTTG CACCTTTGAG CAATGGTCTT TGCTGTGCCT TCTCTACTGA	420
TGGCAGTGTT TTAGCTGCTG GGACACATGA CGGAAGTGTG TATTTTGGG CCACTCCACG	480
GCAGGTCCCT AGCCTGCAAC ATTTATGTGG CATGTCAATC CGAAGAGTGA TGCCCACCCA	540
AGAAGTTCAG GAGCTGCCGA TTCCTTCCAA GCTTTTGGAG TTTCTCTCGT ATCGTATTTA	600
GAAGATTCTG CCTTCCCTAG TAGTAGGGAC TGACAGAATA CACTTAACAC AAACCTCAAG	660
CTTTACTGAC TTCAATTATC TGTTTTAAA GACGTAGAAG ATTTATTTAA TTTGATATGT	720
TCTTGACTG CATTTTGATC AGTTGAGCTT TAAAATATT ATTTATAGAC AATAGAAGTA	780
TTTCTGAACA TATCAAATAT AAATTTTTTT AAAGATCTAA CTGTGAAAAC ATACATACCT	840
GTACATATTT AGATATAAGC TGCTATATGT TGAATGGACC CTTTGTCTTT TCTGATTTTT	900
AGTTCTGACA TGTATATATT GCTTCAGTAG AGCCACAATA TGTATCTTTG CTGTAAAGTG	960
CAAGGAAATT TTAAATTCTG GGACACTGAG TTAGATGGTA AATACTGACT TACGAAAGTT	1020
GAATTGGGTG AGGCGGGCAA ATCACCTGAG GTCAGCAGTT TGAGACTAGC CTGGCAAACA	1080
TGATGAAACC CTGTCTCTAC TAAAAATACA AAAAAAAAAA AA	1122

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2544 base pairs
- (B) TYPE: nucleic acid

B1

- 129 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 423..2030

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGCACGAGC CGGGCTCCGT CCGGAGGAAG CGAGGCTGCG CCGCCGGCCC GGCAGGAGCG	60
GAGGACGGGA MGC GCGGGCG GTCGCGCTCG CCCTGTCGCT GACTGCGCTG CCCC GGCCCA	120
TCCTTGCTG GCCGCAGGTG CCCTGGATGA GCGCGCCGCG CGTGTCCCGG CCGCTGAGTG	180
TCCCCCGCGG TCGCCCGGCG CCTGCCCTCA AGCGGCCGCC TCTCCTTGCC CGGGTCCCCG	240
TTTTCCCCCG GCGCAGTCCT CCTCCGGTGG GCGCCTCCGC ACCTCGGCGC AGGCGGCACG	300
GCCCTCGGGC CGGGATGGAT CCGCCGGGAA GAGGAAGACA AGCCGGGGCG TTGAGCCCCT	360
GCGCACGGTG CCGCCGCGCG TAGTGGGAGC TTA CTGCGAG TAGGCTCTCG CTCTTCTAAT	420
CA ATG GAT AAA GTG GGG AAA ATG TGG AAC AAC TTA AAA TAC AGA TGC	467
Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys	
1 5 10 15	
CAG AAT CTC TTC AGC CAC GAG GGA GGA AGC CGT AAT GAG AAC GTG GAG	515
Gln Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu	
20 25 30	
ATG AAC CCC AAC AGA TGT CCG TCT GTC AAA GAG AAA AGC ATC AGT CTG	563
Met Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu	
35 40 45	
GGA GAG GCA GCT CCC CAG CAA GAG AGC AGT CCC TTA AGA GAA AAT GTT	611
Gly Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val	
50 55 60	
GCC TTA CAG CTG GGA CTG AGC CCT TCC AAG ACC TTT TCC AGG CGG AAC	659
Ala Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn	
65 70 75	
CAA AAC TGT GCC GCA GAG ATC CCT CAA GTG GTT GAA ATC AGC ATC GAG	707
Gln Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu	
80 85 90 95	
AAA GAC AGT GAC TGG GGT GCC ACC CCA GGA ACG AGG CTT GCA CGG AGA	755
Lys Asp Ser Asp Ser Gly Ala Thr Pro Gly Thr Arg Leu Ala Arg Arg	
100 105 110	

B1

- 130 -

GAC	TCC	TAC	TCG	CGG	CAC	GCC	CCG	TGG	GGA	GGA	AAG	AAG	AAA	CAT	TCC	803
Asp	Ser	Tyr	Ser	Arg	His	Ala	Pro	Trp	Gly	Gly	Lys	Lys	Lys	His	Ser	
			115					120					125			
TGT	TCC	ACA	AAG	ACC	CAG	AGT	TCA	TTG	GAT	ACC	GAG	AAA	AAG	TTT	GGT	851
Cys	Ser	Thr	Lys	Thr	Gln	Ser	Ser	Leu	Asp	Thr	Glu	Lys	Lys	Phe	Gly	
		130					135					140				
AGA	ACT	CGA	AGC	GGC	CTT	CAG	AGG	CGA	GAG	CGG	CGC	TAT	GGA	GTC	AGC	899
Arg	Thr	Arg	Ser	Gly	Leu	Gln	Arg	Arg	Glu	Arg	Arg	Tyr	Gly	Val	Ser	
	145					150					155					
TCC	ATG	CAG	GAC	ATG	GAC	AGC	GTT	TCT	AGC	CGC	GCG	GTC	GGG	AGC	CGC	947
Ser	Met	Gln	Asp	Met	Asp	Ser	Val	Ser	Ser	Arg	Ala	Val	Gly	Ser	Arg	
	160				165					170					175	
TCC	CTG	AGG	CAG	AGG	CTC	CAG	GAC	ACG	GTG	GGT	TTG	TGT	TTT	CCC	ATG	995
Ser	Leu	Arg	Gln	Arg	Leu	Gln	Asp	Thr	Val	Gly	Leu	Cys	Phe	Pro	Met	
				180					185					190		
AGA	ACT	TAC	AGC	AAG	CAG	TCA	AAG	CCA	CTC	TTT	TCC	AAT	AAA	AGA	AAA	1043
Arg	Thr	Tyr	Ser	Lys	Gln	Ser	Lys	Pro	Leu	Phe	Ser	Asn	Lys	Arg	Lys	
			195					200					205			
ATA	CAT	CTT	TCT	GAA	TTA	ATG	CTG	GAG	AAA	TGC	CCT	TTT	CCT	GCT	GGC	1091
Ile	His	Leu	Ser	Glu	Leu	Met	Leu	Glu	Lys	Cys	Pro	Phe	Pro	Ala	Gly	
		210					215					220				
TCG	GAT	TTA	GCA	CAA	AAG	TGG	CAT	TTG	ATT	AAA	CAG	CAT	ACC	GCC	CCT	1139
Ser	Asp	Leu	Ala	Gln	Lys	Trp	His	Leu	Ile	Lys	Gln	His	Thr	Ala	Pro	
	225					230					235					
GTG	AGC	CCA	CAC	TCA	ACA	TTT	TTT	GAT	ACA	TTT	GAT	CCA	TCA	CTG	GTG	1187
Val	Ser	Pro	His	Ser	Thr	Phe	Phe	Asp	Thr	Phe	Asp	Pro	Ser	Leu	Val	
	240				245				250					255		
TCT	ACA	GAA	GAT	GAA	GAA	GAT	AGG	CTT	CGC	GAG	AGA	AGA	CGG	CTT	AGT	1235
Ser	Thr	Glu	Asp	Glu	Glu	Asp	Arg	Leu	Arg	Glu	Arg	Arg	Arg	Leu	Ser	
				260					265					270		
ATC	GAA	GAA	GGG	GTG	GAT	CCC	CCT	CCC	AAC	GCA	CAA	ATA	CAC	ACC	TTT	1283
Ile	Glu	Glu	Gly	Val	Asp	Pro	Pro	Pro	Asn	Ala	Gln	Ile	His	Thr	Phe	
			275					280					285			
GAA	GCT	ACT	GCA	CAG	GTC	AAC	CCA	TTG	TAT	AAG	CTG	GGA	CCA	AAG	TTA	1331
Glu	Ala	Thr	Ala	Gln	Val	Asn	Pro	Leu	Tyr	Lys	Leu	Gly	Pro	Lys	Leu	
		290					295					300				
GCT	CCT	GGG	ATG	ACA	GAG	ATA	AGT	GGA	GAT	GGT	TCT	GCA	ATT	CCA	CAA	1379
Ala	Pro	Gly	Met	Thr	Glu	Ile	Ser	Gly	Asp	Gly	Ser	Ala	Ile	Pro	Gln	
	305					310					315					



GCS Xaa 320	AAT Asn	TGT Cys	GAC Asp	TCA Ser	GAA Glu 325	GAG Glu	GAT Asp	TCA Ser	ACC Thr	ACC Thr 330	CTA Leu	TGT Cys	CTG Leu	CAG Gln	TCA Ser 335	1427
CGG Arg	AGG Arg	CAG Gln	AAG Lys	CAG Gln 340	CGC Arg	CAG Gln	GTG Val	TCC Ser	GGG Gly 345	GAC Asp	AGC Ser	CAC His	GCG Ala	CAC His 350	GTT Val	1475
AGC Ser	AGA Arg	CAG Gln	GGA Gly 355	GCT Ala	TGG Trp	AAA Lys	GTT Val	CAT His 360	ACG Thr	CAG Gln	ATC Ile	GAT Asp	TAC Tyr 365	ATA Ile	CAC His	1523
TGC Cys	CTC Leu	GTG Val 370	CCA Pro	GAT Asp	TTG Leu	CTT Leu	CAG Gln 375	ATC Ile	ACA Thr	GGG Gly	AAT Asn 380	CCC Pro	TGT Cys	TAC Tyr	TGG Trp	1571
GGC Gly	GTG Val 385	ATG Met	GAC Asp	CGA Arg	TAC Tyr	GAG Glu 390	GCC Ala	GAA Glu	GCC Ala	CTT Leu	CTA Leu 395	GAA Glu	GGG Gly	AAA Lys	CCG Pro	1619
GAA Glu 400	GGC Gly	ACG Thr	TTC Phe	TTG Leu	CTC Leu 405	AGG Arg	GAC Asp	TCT Ser	GCA Ala	CAG Gln 410	GAG Glu	GAC Asp	TAC Tyr	CTC Leu	TTC Phe 415	1667
TCT Ser	GTG Val	AGC Ser	TTC Phe	CGC Arg 420	CGC Arg	TAC Tyr	AAC Asn	AGG Arg	TCT Ser	CTG Leu	CAC His	GCC Ala	CGG Arg	ATC Ile 430	GAG Glu	1715
CAG Gln	TGG Trp	AAC Asn 435	CAC His	AAC Asn	TTC Phe	AGC Ser	TTC Phe	GAT Asp 440	GCC Ala	CAT His	GAC Asp	CCC Pro	TGC Cys 445	GTG Val	TTT Phe	1763
CAC His	TCC Ser	TCC Ser 450	ACW Xaa	GTC Val	ACG Thr	GGG Gly	CTT Leu 455	CTC Leu	GAA Glu	CAC His	TAT Tyr	AAA Lys 460	GAC Asp	CCC Pro	AGC Ser	1811
TCT Ser	TGC Cys 465	ATG Met	TTT Phe	TTT Phe	GAA Glu	CCG Pro 470	TTG Leu	CTA Leu	ACG Thr	ATA Ile	TCA Ser	CTG Leu	AAT Asn	AGA Arg	ACT Thr	1859
TTC Phe 480	CCT Pro	TTC Phe	AGC Ser	CTG Leu	CAG Gln 485	TAT Tyr	ATC Ile	TGC Cys	CGC Arg	GCA Ala 490	GTG Val	ATC Ile	TGC Cys	AGA Arg	TGC Cys 495	1907
ACT Thr	ACG Thr	TAT Tyr	GAT Asp	GGG Gly 500	ATT Ile	GAC Asp	GGG Gly	CTC Leu	CCG Pro 505	CTA Leu	CCG Pro	TCG Ser	ATG Met	TTA Leu 510	CAG Gln	1955
GAT Asp	TTT Phe	TTA Leu	AAA Lys 515	GAG Glu	TAT Tyr	CAT His	TAT Tyr	AAA Lys 520	CAA Gln	AAA Lys	GTT Val	AGG Arg	GTT Val	CGC Arg	TGG Trp	2003

*Handwritten signature/initials*

- 132 -

TTA GAA CGA GAR CCA GTC AAA GCA AAG TAACTCCTGT CCCCAAAGGG 2050  
Leu Glu Arg Xaa Pro Val Lys Ala Lys  
530 535

CACTAACTAA GTCTGCTCCT CCCGTGCATC MGAAGTGCAC CCATAGGRAG GCAGTCAGCT 2110  
GCTAGGATTT CCCACCCAGA ATGGGAGCTT AGTCATTAGC CTCTGCCCTA TGGGGTCCGC 2170  
TGTTCTCCTCAG ACAAAGGTGC CTAGGGACAG CAAGATGGCT TGCAGGTGTT CGGTGGGCTG 2230  
TGACAACTGA GGGAGGCAAC TCTGGGGCAT TTGCTATGAA GAATTCTATT TCTTACCGAA 2290  
GAACAAATTA TTAATATTGG ATGGGTATTT CAATAGTGTG ACTAATGTTT GAAATTATTT 2350  
TTTCTAAGAA TTTTCTATA ACCTTCAGAA AAAGTAGTGA TGTTTGTAGT TACTATAAAT 2410  
CAAGCTTTGA AAGTTCAAAA CAAACAAGTT AAATAAAAGA CTACCTTCCT TTTAGAGAAA 2470  
ACAAATGCAA GTTTTCCCAG CCACAGGCAT TGTGCACTGT TAATGTTAGC TTGTTATCAG 2530  
CTCCTTTCTC CTCC 2544

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 536 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Asp Lys Val Gly Lys Met Trp Asn Asn Leu Lys Tyr Arg Cys Gln  
1 5 10 15  
Asn Leu Phe Ser His Glu Gly Gly Ser Arg Asn Glu Asn Val Glu Met  
20 25 30  
Asn Pro Asn Arg Cys Pro Ser Val Lys Glu Lys Ser Ile Ser Leu Gly  
35 40 45  
Glu Ala Ala Pro Gln Gln Glu Ser Ser Pro Leu Arg Glu Asn Val Ala  
50 55 60  
Leu Gln Leu Gly Leu Ser Pro Ser Lys Thr Phe Ser Arg Arg Asn Gln  
65 70 75 80  
Asn Cys Ala Ala Glu Ile Pro Gln Val Val Glu Ile Ser Ile Glu Lys  
85 90 95

*Bl cont*

- 133 -

Asp	Ser	Asp	Ser	Gly	Ala	Thr	Pro	Gly	Thr	Arg	Leu	Ala	Arg	Arg	Asp
			100					105					110		
Ser	Tyr	Ser	Arg	His	Ala	Pro	Trp	Gly	Gly	Lys	Lys	Lys	His	Ser	Cys
		115					120					125			
Ser	Thr	Lys	Thr	Gln	Ser	Ser	Leu	Asp	Thr	Glu	Lys	Lys	Phe	Gly	Arg
	130					135					140				
Thr	Arg	Ser	Gly	Leu	Gln	Arg	Arg	Glu	Arg	Arg	Tyr	Gly	Val	Ser	Ser
145					150					155					160
Met	Gln	Asp	Met	Asp	Ser	Val	Ser	Ser	Arg	Ala	Val	Gly	Ser	Arg	Ser
				165					170					175	
Leu	Arg	Gln	Arg	Leu	Gln	Asp	Thr	Val	Gly	Leu	Cys	Phe	Pro	Met	Arg
			180					185					190		
Thr	Tyr	Ser	Lys	Gln	Ser	Lys	Pro	Leu	Phe	Ser	Asn	Lys	Arg	Lys	Ile
		195					200					205			
His	Leu	Ser	Glu	Leu	Met	Leu	Glu	Lys	Cys	Pro	Phe	Pro	Ala	Gly	Ser
	210					215					220				
Asp	Leu	Ala	Gln	Lys	Trp	His	Leu	Ile	Lys	Gln	His	Thr	Ala	Pro	Val
225					230					235					240
Ser	Pro	His	Ser	Thr	Phe	Phe	Asp	Thr	Phe	Asp	Pro	Ser	Leu	Val	Ser
				245					250					255	
Thr	Glu	Asp	Glu	Glu	Asp	Arg	Leu	Arg	Glu	Arg	Arg	Arg	Leu	Ser	Ile
			260					265					270		
Glu	Glu	Gly	Val	Asp	Pro	Pro	Pro	Asn	Ala	Gln	Ile	His	Thr	Phe	Glu
		275					280					285			
Ala	Thr	Ala	Gln	Val	Asn	Pro	Leu	Tyr	Lys	Leu	Gly	Pro	Lys	Leu	Ala
	290					295					300				
Pro	Gly	Met	Thr	Glu	Ile	Ser	Gly	Asp	Gly	Ser	Ala	Ile	Pro	Gln	Xaa
305					310					315					320
Asn	Cys	Asp	Ser	Glu	Glu	Asp	Ser	Thr	Thr	Leu	Cys	Leu	Gln	Ser	Arg
				325					330					335	
Arg	Gln	Lys	Gln	Arg	Gln	Val	Ser	Gly	Asp	Ser	His	Ala	His	Val	Ser
			340					345					350		
Arg	Gln	Gly	Ala	Trp	Lys	Val	His	Thr	Gln	Ile	Asp	Tyr	Ile	His	Cys
		355					360					365			

Leu Val Pro Asp Leu Leu Gln Ile Thr Gly Asn Pro Cys Tyr Trp Gly  
 370 375 380  
 Val Met Asp Arg Tyr Glu Ala Glu Ala Leu Leu Gln Gly Lys Pro Glu  
 385 390 395 400  
 Gly Thr Phe Leu Leu Arg Asp Ser Ala Gln Glu Asp Tyr Leu Phe Ser  
 405 410 415  
 Val Ser Phe Arg Arg Tyr Asn Arg Ser Leu His Ala Arg Ile Glu Gln  
 420 425 430  
 Trp Asn His Asn Phe Ser Phe Asp Ala His Asp Pro Cys Val Phe His  
 435 440 445  
 Ser Ser Xaa Val Thr Gly Leu Leu Glu His Tyr Lys Asp Pro Ser Ser  
 450 455 460  
 Cys Met Phe Phe Glu Pro Leu Leu Thr Ile Ser Leu Asn Arg Thr Phe  
 465 470 475 480  
 Pro Phe Ser Leu Gln Tyr Ile Cys Arg Ala Val Ile Cys Arg Cys Thr  
 485 490 495  
 Thr Tyr Asp Gly Ile Asp Gly Leu Pro Leu Pro Ser Met Leu Gln Asp  
 500 505 510  
 Phe Leu Lys Glu Tyr His Tyr Lys Gln Lys Val Arg Val Arg Trp Leu  
 515 520 525  
 Glu Arg Xaa Pro Val Lys Ala Lys  
 530 535

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1221 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATTAAACAG CATAACGCTC CTGTGAGCCC ACATTCAACA TTTTGTGATA CTTTGATCCA	60
TCTTTGGTTT CTACAGAAGA TGAAGAAGAT AGGCTTAGAG AGAGAAGGCG GCTTAGTATT	120
GAAGAAGGGG TTGATCCCCC TCCAATGCA CAAATACATA CATTTGAAGC TACTGCACAG	180

B1

- 135 -

GTTAATCCAT TATTAACTG GGACCAAAAT TAGCTCCTGG AATGACTGAA ATAAGTGGGG	240
ACAGTTCTGC AATTCCACAA GCTAATTGTG ACTCGGAAGA GGATACAACC ACCCTGTGTT	300
GCAGTCACGG AGGCAGAAGC AGCGTCAGAT ATCTGGAGAC AGGCATACCC ATGTTAGCAG	360
ACAGGGAGCT TGGAAAGTCC ACACACAGAT TGATTACATA CACTGCTTCG TGCCTGATTT	420
GCTTCAAATT ACAGGGAATC CCTGTTACTG GGGAGTGATG GACCGTTATG AAGCAGAAGC	480
CCTTCTCGAA GGGAAACCTG AAGGCACGTT TTTGCTCAGG GACTCTGCGC AAGAGGACTA	540
CTTCTTCTCT GTGAGCTTCC GCCGATACAA CAGATCCTCG CATGCCCCGAA TTGAGCAGTG	600
GAATCACAAC TTTAGTTTCG ACGCCCATGA CCCGTTGTGTA TTTCACCTCCT CCACTGTAAC	660
GGGACTTTTA GAACATTATA AAGATCCCAG TTCCTGTCATG TTTTTTGAAC CATTGCTTAC	720
TATATCACTA AATAGGACTT TCCCTTTTAG CCGTGCAGTAT ATCTGTCGCG CGGTAATCTG	780
CAGGTGCACT ACGTATGATG GAATTGATGG GCTCCCTCTA CCCTCAATGT TACAGGATTT	840
TTTAAAAGAG TATCATTATA AACAAAAAGT TAGAGTTCGC TGGTTGGAAC GAGAACCAGT	900
CAAGGCAAAG TAAACTCTCC GGTCCCCAAA GGGTGTTAAC TAGGTCCGCT TTCATGTGCA	960
TCAGACAGTA CACCTATAGC AAGCACACGT AGCAGTGTTA GGCTTTTTTCA TACAGTATGT	1020
AAGCTTAGTG TTAGTATCTG TCAGATGCTA CCTGCTGTTA CTTATTCAGA TAAACATGGT	1080
GCCTATTGGA ACAATAGCGG ATAGAGCTAC AGGTGTTTCA TAAGACTACA AAAACATTTT	1140
GCCTATTTTG CTAACAGTTT GGTTTTTTAAT GGCTGTGGTA TTTGAGTGAG GCAACTCTGG	1200
GGCATTGTGTT ATGAAGAAAT G	1221

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2369 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 116..1330

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

B1

- 136 -

GGCACGAGGC	GGTGGTGGCG	GCGGCGGGCG	CGGCCGCGGC	GGGGCGGGCG	CGGAATGAAG	60
GCCACGGCC	CTGGGGGCTG	AGGCGCCCGC	CGCCTGGGGC	GGGCCGCGCG	TCCTC ATG	118
					Met	
					1	
GAG GCC GGA GAG GAG CCG CTG CTG CTG GCT GAA CTC AAG CCT GGG CGC	166					
Glu Ala Gly Glu Glu Pro Leu Leu Leu Ala Glu Leu Lys Pro Gly Arg						
5 10 15						
CCC CAC CAG TTC GAC TGG AAG TCA AGC TGC GAG ACC TGG AGC GTG GCC	214					
Pro His Gln Phe Asp Trp Lys Ser Ser Cys Glu Thr Trp Ser Val Ala						
20 25 30						
TTC TCG CCA GAC GGT TCC TGG TTC GCC TGG TCT CAA GGA CAC TGC GTG	262					
Phe Ser Pro Asp Gly Ser Trp Phe Ala Trp Ser Gln Gly His Cys Val						
35 40 45						
GTC AAG CTG GTC CCC TGG CCC TTA GAG GAA CAG TTC ATC CCT AAA GGA	310					
Val Lys Leu Val Pro Trp Pro Leu Glu Glu Gln Phe Ile Pro Lys Gly						
50 55 60 65						
TTC GAA GCC AAG AGC CGA AGC AGC AAG AAT GAC CCA AAA GGA CGG GGC	358					
Phe Glu Ala Lys Ser Arg Ser Ser Lys Asn Asp Pro Lys Gly Arg Gly						
70 75 80						
AGT CTG AAG GAG AAG ACG CTG GAC TGT GGC CAG ATT GTG TGG GGG CTG	406					
Ser Leu Lys Glu Lys Thr Leu Asp Cys Gly Gln Ile Val Trp Gly Leu						
85 90 95						
GCC TTC AGC CCG TGG CCC TCT CCA CCC AGC AGG AAA CTC TGG GCA CGT	454					
Ala Phe Ser Pro Trp Pro Ser Pro Pro Ser Arg Lys Leu Trp Ala Arg						
100 105 110						
CAC CAT CCC CAG GCG CCT GAT GTT TCT TGC CTG ATC CTG GCC ACA GGT	502					
His His Pro Gln Ala Pro Asp Val Ser Cys Leu Ile Leu Ala Thr Gly						
115 120 125						
CTC AAC GAT GGG CAG ATC AAG ATT TGG GAG GTA CAG ACA GGC CTC CTG	550					
Leu Asn Asp Gly Gln Ile Lys Ile Trp Glu Val Gln Thr Gly Leu Leu						
130 135 140 145						
CTT CTG AAT CTT TCT GGC CAC CAA GAC GTC GTG AGA GAT CTG AGC TTC	598					
Leu Leu Asn Leu Ser Gly His Gln Asp Val Val Arg Asp Leu Ser Phe						
150 155 160						
ACG CCC AGC GGC AGT TTG ATT TTG GTC TCT GCA TCC CGG GAT AAG ACA	646					
Thr Pro Ser Gly Ser Leu Ile Leu Val Ser Ala Ser Arg Asp Lys Thr						
165 170 175						

CTT	CGA	ATT	TGG	GAC	CTG	AAT	AAA	CAC	GGT	AAG	CAG	ATC	CAG	GTG	TTA	694
Leu	Arg	Ile	Trp	Asp	Leu	Asn	Lys	His	Gly	Lys	Gln	Ile	Gln	Val	Leu	
	180						185					190				
TCC	GGC	CAT	CTG	CAG	TGG	GTT	TAC	TGC	TGC	TCC	ATC	TCC	CCT	GAC	TGT	742
Ser	Gly	His	Leu	Gln	Trp	Val	Tyr	Cys	Cys	Ser	Ile	Ser	Pro	Asp	Cys	
	195					200					205					
AGC	ATG	CTG	TGC	TCT	GCA	GCT	GGG	GAG	AAG	TGG	GTC	TTT	CTG	TGG	AGC	790
Ser	Met	Leu	Cys	Ser	Ala	Ala	Gly	Glu	Lys	Ser	Val	Phe	Leu	Trp	Ser	
210					215					220					225	
ATG	CGG	TCC	TAC	ACA	CTA	ATC	CGG	AAA	CTA	GAA	GGC	CAC	CAA	AGC	AGT	838
Met	Arg	Ser	Tyr	Thr	Leu	Ile	Arg	Lys	Leu	Glu	Gly	His	Gln	Ser	Ser	
				230					235					240		
GTT	GTC	TCC	TGT	GAT	TTC	TCT	CCT	GAT	TCA	GCC	TTG	CTT	GTC	ACA	GCT	886
Val	Val	Ser	Cys	Asp	Phe	Ser	Pro	Asp	Ser	Ala	Leu	Leu	Val	Thr	Ala	
			245					250					255			
TCG	TAT	GAC	ACC	AGT	GTG	ATT	ATG	TGG	GAC	CCC	TAC	ACC	GGC	GCG	AGG	934
Ser	Tyr	Asp	Thr	Ser	Val	Ile	Met	Trp	Asp	Pro	Tyr	Thr	Gly	Ala	Arg	
	260						265					270				
CTG	AGG	TCA	CTT	CAT	CAC	ACA	CAA	CTT	GAA	CCC	ACC	ATG	GAT	GAC	AGT	982
Leu	Arg	Ser	Leu	His	His	Thr	Gln	Leu	Glu	Pro	Thr	Met	Asp	Asp	Ser	
	275					280					285					
GAC	GTC	CAC	ATG	AGC	TCC	CTG	AGG	TCC	GTG	TGC	TTC	TCA	CCT	GAA	GGC	1030
Asp	Val	His	Met	Ser	Ser	Leu	Arg	Ser	Val	Cys	Phe	Ser	Pro	Glu	Gly	
290					295					300					305	
TTG	TAT	CTC	GCT	ACG	GTG	GCA	GAT	GAC	AGG	CTG	CTC	AGG	ATC	TGG	GCT	1078
Leu	Tyr	Leu	Ala	Thr	Val	Ala	Asp	Asp	Arg	Leu	Leu	Arg	Ile	Trp	Ala	
				310					315					320		
CTG	GAA	CTG	AAG	GCT	CCG	GTT	GCC	TTT	GCT	CCG	ATG	ACC	AAT	GGT	CTT	1126
Leu	Glu	Leu	Lys	Ala	Pro	Val	Ala	Phe	Ala	Pro	Met	Thr	Asn	Gly	Leu	
			325					330					335			
TGC	TGC	ACG	TTC	TTC	CCA	CAC	GGT	GGA	ATT	ATT	GCC	ACA	GGG	ACG	AGA	1174
Cys	Cys	Thr	Phe	Phe	Pro	His	Gly	Gly	Ile	Ile	Ala	Thr	Gly	Thr	Arg	
		340					345					350				
GAT	GGC	CAT	GTC	CAG	TTC	TGG	ACA	GCT	CCC	CGG	GTC	CTG	TCC	TCA	CTG	1222
Asp	Gly	His	Val	Gln	Phe	Trp	Thr	Ala	Pro	Arg	Val	Leu	Ser	Ser	Leu	
	355					360					365					
AAG	CAC	TTA	TGC	AGG	AAA	GCC	CTC	CGA	AGT	TTC	CTG	ACA	ACG	TAT	CAA	1270
Lys	His	Leu	Cys	Arg	Lys	Ala	Leu	Arg	Ser	Phe	Leu	Thr	Thr	Tyr	Gln	
370					375					380					385	

GTC CTA GCA CTG CCA ATC CCC AAG AAG ATG AAA GAG TTC CTC ACA TAC	1318
Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr Tyr	
390 395 400	
AGG ACT TTC TAGCAGTGCC GGCTCCCCCA CCTCCTGCAG CAGCAGCAGT	1367
Arg Thr Phe	
405	
ACAAGGGACT GGCTAGGATG GAGTCAGGCA GCTCACACTG GACCAGTGTG GACCTTCCTT	1427
CCTCCCATGG CATGTGCAAG TAGGTCTGCG TGACCCCACT TCTGTGGTGC CGGCCTTACC	1487
TCGTCTTCAT CCGTGGTGAG CAGCCTTCGT CAGTCTAGTT GTGTTGAAGC CAAGTGCAGT	1547
TGTGGATGTT GCTGGGGTAA TAAAGGCAAG CGGGCTCCAG AGCCTCTCTG GTGGCGGCCA	1607
AGCCACACTC CCTTAACTGG GAAGTACCTG CCAAGTAGGG CATTTCTGCT GCCTATTTCC	1667
AGCCAGCGGC TGCATGGTTT GAAGTTCCTC CGTTGTGGTC AGAAGAACTC TGGTGTTTGG	1727
TTCCCTGCTC AGCTGCGCGT GGACTGGGCT GAGCTCCTCA CCATACACTA GTGCCGGCTT	1787
TTGTTTCCTG TAAACAGTGG TTGCATGTGT AGAGAAGTAA CAAGCGAGTA TTCAGATCAT	1847
ACGAGGAGGC GTTCCTCGGT GCATGACGGT CAGATGGCCA TTTATCAGCA TATTTATTTG	1907
TATTTTCTCA GCACATAGTA AGGTACAAC GTGTTTTCTC AATTGTCTCG AAAAAACAGA	1967
GTTCTTAAGT GGCCCAGTTG TGGAGCOAAG TCTAAGTCGT GTGGAGTCAG TGCTGACATC	2027
ACTGGCTTGT GCTGTCTGTC ACATGTGTTT GTCTCTGCTG CTTGACCTCA TGGGATGTAC	2087
CCTCCAGTTC AACTGCCCAA AACAGACAGC CCCTTCCAAG CACCGTTCTT TGACAGCGGT	2147
AGCAGCTACC TATTCAAGAC GCCTCACACA AAATCTGCCT TAGAAAGTTA ATATATTTTA	2207
AATTATTTTA AAAGAACTC AACATCTTAT TCTTTGGCCT TTCTTAATTG ATGCTTTATG	2267
GAGGCAGTGT TAACATTGTA CAGTGTATGC ATAGAGGAGT CTCCTCTATT TGAAGAACAA	2327
TGCAAAATGA GGCTTTCATT GAAGGGAAAA AAAAAAAAAA AA	2369

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 404 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:



*Al  
Conf*

- 139 -

Met	Glu	Ala	Gly	Glu	Glu	Pro	Leu	Leu	Leu	Ala	Glu	Leu	Lys	Pro	Gly
1				5				10						15	
Arg	Pro	His	Gln	Phe	Asp	Trp	Lys	Ser	Ser	Cys	Glu	Thr	Trp	Ser	Val
			20					25					30		
Ala	Phe	Ser	Pro	Asp	Gly	Ser	Trp	Phe	Ala	Trp	Ser	Gln	Gly	His	Cys
		35					40					45			
Val	Val	Lys	Leu	Val	Pro	Trp	Pro	Leu	Glu	Glu	Gln	Phe	Ile	Pro	Lys
	50					55					60				
Gly	Phe	Glu	Ala	Lys	Ser	Arg	Ser	Ser	Lys	Asn	Asp	Pro	Lys	Gly	Arg
65					70					75					80
Gly	Ser	Leu	Lys	Glu	Lys	Thr	Leu	Asp	Cys	Gly	Gln	Ile	Val	Trp	Gly
			85						90					95	
Leu	Ala	Phe	Ser	Pro	Trp	Pro	Ser	Pro	Pro	Ser	Arg	Lys	Leu	Trp	Ala
			100					105					110		
Arg	His	His	Pro	Gln	Ala	Pro	Asp	Val	Ser	Cys	Leu	Ile	Leu	Ala	Thr
		115					120					125			
Gly	Leu	Asn	Asp	Gly	Gln	Ile	Lys	Ile	Trp	Glu	Val	Gln	Thr	Gly	Leu
130						135					140				
Leu	Leu	Leu	Asn	Leu	Ser	Gly	His	Gln	Asp	Val	Val	Arg	Asp	Leu	Ser
145					150					155					160
Phe	Thr	Pro	Ser	Gly	Ser	Leu	Ile	Leu	Val	Ser	Ala	Ser	Arg	Asp	Lys
			165						170					175	
Thr	Leu	Arg	Ile	Trp	Asp	Leu	Asn	Lys	His	Gly	Lys	Gln	Ile	Gln	Val
		180						185					190		
Leu	Ser	Gly	His	Leu	Gln	Trp	Val	Tyr	Cys	Cys	Ser	Ile	Ser	Pro	Asp
		195					200					205			
Cys	Ser	Met	Leu	Cys	Ser	Ala	Ala	Gly	Glu	Lys	Ser	Val	Phe	Leu	Trp
210						215					220				
Ser	Met	Arg	Ser	Tyr	Thr	Leu	Ile	Arg	Lys	Leu	Glu	Gly	His	Gln	Ser
225					230					235					240
Ser	Val	Val	Ser	Cys	Asp	Phe	Ser	Pro	Asp	Ser	Ala	Leu	Leu	Val	Thr
			245						250					255	
Ala	Ser	Tyr	Asp	Thr	Ser	Val	Ile	Met	Trp	Asp	Pro	Tyr	Thr	Gly	Ala
			260					265					270		

Arg Leu Arg Ser Leu His His Thr Gln Leu Glu Pro Thr Met Asp Asp  
 275 280 285  
 Ser Asp Val His Met Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu  
 290 295 300  
 Gly Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp  
 305 310 315 320  
 Ala Leu Glu Leu Lys Ala Pro Val Ala Phe Ala Pro Met Thr Asn Gly  
 325 330 335  
 Leu Cys Cys Thr Phe Phe Pro His Gly Gly Ile Ile Ala Thr Gly Thr  
 340 345 350  
 Arg Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser  
 355 360 365  
 Leu Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr  
 370 375 380  
 Gln Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr  
 385 390 395 400  
 Tyr Arg Thr Phe

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAACTGCAT CGTCAAACCTG ATCCCCTGGC CGTTGGAGGA GCAGTTCATC CCTAAAGGGT	60
TTGAAGCCAA AAGCCGAAGT AGCAAAAATG AGACGAAAGG GCGGGGCAGC CCAAAAGAGA	120
AGACGCTGGA CTGTGGTCAG ATTGTCTGGG GGCTGGCCTT CAGCCTGTGC TTTCCCCACC	180
CAGCAGGAAG CTCTGGGCAC GCCACCACCC CCAAGTGCCC GATGTCTCTT GCCTGGTTCT	240
TGCTACGGGA CTCAACGATG GGCAGATCAA GATCTGGGAG GTGCAGACAG GGCTCCTGCT	300
TTTGAATCTT TCCGGCCACC AAGATGTCGT GAGAGATCTG AGCTTCACAC CCAGTGGCAG	360
TTTGATTTTG GTCTCCGCGT CACGGGATAA GACTCTTCGC ATCTGGGACC TGAATAAACA	420

CGGTAAACAG ATTCAAGTGT TATCGGGCCA CCTGCAGTGG GTTTACTGCT GTTCCATCTC	480
CCCAGACTGC AGCATGCTGT GCTCTGCAGC TGGAGAGAAG TCGGTCTTTC TATGGAGCAT	540
GAGGTCCTAC ACGTTAATTC GGAAGCTAGA GGGCCATCAA AGCAGTGTTG TCTCTTGTA	600
CTTCTCCCC GACTCTGCCC TGCTTGTCAC GGCTTCTTAC GATACCAATG TGATTATGTG	660
GGACCCCTAC ACCGGCGAAA GGCTGAGGTC ACTCCACCAC ACCCAGGTTG ACCCCGCCAT	720
GGATGACAGT GACGTCCACA TTAGCTCACT GAGATCTGTG TGCTTCTCTC CAGAAGGCTT	780
GTACCTTGCC ACGGTGGCAG ATGACAGACT CCTCAGGATC TGGGCCCTGG AACTGAAAAC	840
TCCCATTGCA TTTGCTCCTA TGACCAATGG GCTTTGOTGG CACATTTTTT CCACATGGTG	900
GAGTCATTGC CACAGGGACA AGAGATGGCC ACGTCCAGTT CTGGACAGCT CCTAGGGTCC	960
TGTCCTCACT GAAGCACTTA TGCCGGAAAG CCCTTCGAAG TTTCCTAACA ACTTACCAAG	1020
TCCTAGCACT GCCAATCCCC AAGAAAATGA AAGAGTTCCT CACATACAGG ACTTTTAAAG	1080
CAACACCACA TCTTGTGCTT CTTTGTAGCA GGGTAAATCG TCCTGTCAA GGGAGTTGCT	1140
GGAATAATGG GCCAAACATC TGGTCTTGCA TTGAAATAGC ATTTCTTTGG GATTGTGAAT	1200
AGAATGTAGC AAAACCAGAT TCCAGTGTAC TAGTCATGGA TTTTTC	1246

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 422 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ACCATGGTTC CAAGTCCTCT CCCCTGTGGT CAAGTTGCCC GAATGTTGGG CCCAAGTGCC	60
TTTTCCTCCT TGGGCCTCCC CTTGTGACCT GCAGGACAGT TTTCCGGAGC CCATTGCGTA	120
TGAGGTATTA ATTAGCCTTA ACTAAATTAC AGGGGACTCA GAGGCCGTGC TCCTGACCGA	180
TCCAGACACT ATTTTTTTTTT TTTTTTTTTTA ACAATGGTGT GCATGTGCAG GAAATGACAA	240
ATTTGTATGT CAGATTATAC AAGGATGTAT TCTTAAACCG CATGACTATT CAGATGGCTA	300
CTGAGTTATC AGTGGCCATT TATTAGCATC ATATTTATTT GTATTTTCTC AACAGATGTT	360

AAGGTACAAC TGTGTTTTTC TCGATTATCT AAAAACCATA GTACTTAAAT TGAAAAAAAA 420

AA 422

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2019 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGCACGAGGC GGGGTCAGGG CGGAGGCTGA GGACCAAGTA GGCATGGCGG AGGGCGGGAC 60  
CGGCCCCGAT GGACGGGCCG GCCCGGGACC CGCAGGTCCT AATCTGAAGG AGTGGCTGAG 120  
GGAGCAGTTC TGTGACCATC CACTGGAGCA CTGTGACGAT ACAAGACTCC ATGATGCAGC 180  
CTATGTAGGG GACCTCCAGA CCCTCAGGAA CCTACTGCAA GAGGAGAGCT ACCGGAGCCG 240  
CATCAATGAG AAGTCTGTCT GGTGCTGCGG CTGGCTTCCC TGCACACCAC TGAGGATCGC 300  
AGCCACTGCA GGCCATGGGA ACTGTGTGGA CTTCCTCATA CGCAAAGGGG CCGAGGTGGA 360  
CCTGGTGGAT GTCAAGGGGC AGACTGCCCT GTATGTGGCT GTAGTGAACG GGCACCTGGA 420  
GAGCACTGAG ATCCTTTTGG AAGCTGGTGC TGATCCCAAC GGCAGCCGGC ACCACCGCAG 480  
CACTCCTGTG TACCATGCCT YTCGTGTGGG TAGGGACGAC ATCCTGAAGG CTCTTATCAG 540  
GTATGGGGCA GATGTTGATG TCAACCATCA TCTGAATTCT GACACCCGGC CCCCTTTTTC 600  
ACGGCGGCTA ACCTCCTTGG TGGTCTGTCC TCTATACATC AGTGCTGCCT ACCATAACCT 660  
TCAGTGCTTC AGGCTGCTCT TGCAAGCTGG GGCAAATCCT GACTTCAATT GCAATGGCCC 720  
TGTCACACC CAGGAGTTCT ACAGGGGATC CCCTGGGTGT GTCATGGATG CTGTCCTGCG 780  
CCATGGCTGT GAAGCAGCCT TCGTGAGTCT GTTGGTAGAG TTTGGAGCCA ACCTGAACCT 840  
GGTGAAGTGG GAATCCCTGG GCCCAGAGGC AAGAGGCAGA AGAAAGATGG ATCCTGAGGC 900  
CTTGCAAGTC TTAAAGAGG CCAGAAGTAT TCCCAGGACC TTGCTGAGTT TGTGCCGGGT 960  
GGCTGTGAGA AGAGCTCTTG GCAAATACCG ACTGCATCTG GTTCCCTCGC TGCCGCTGCC 1020  
AGACCCATA AAGAAGTTT TGCTTTATGA GTAGCATTCA CATGCAGTGC TGA CTGCAAT 1080

BL  
cont

GTGGAAGCCG ATCACCTGCA GTGAAACTG ACACAGACTC TGGCATCCTG GGAACCATGG 1140  
 CCTGTGCTGC CAGCTTGATC CTTGGCTGTC AGTGAAGAAA AAACGGCTGT GTTCTCTTGG 1200  
 ACTGTGATTC TATCTCAGGT GCTTGGGCCA TCGAACGCTC CTTGAGTCAT TGTCAACTGA 1260  
 GAGGCACATA CAAACTTAAT TTTGTTCTC TTCAGTCTCT CTGTTTGGGA TTCTTCCTGG 1320  
 CAATGTGTGC AGCATGGGCT GAGCCTGGTG ATTGCCCTAG TGGGGAAGGC TTTTCTCTCC 1380  
 AGGCTATGCA TCTATTTATG TTCCTACTTT GCAATTTATT GTTCTTTTAA GGCTTGATAT 1440  
 CAAAACAGAA AGAGGTTTGT TAAGAAAAGA TATAGGGAGA AAGGAATTCC GGTTCCTGTC 1500  
 ACTTGCTAGC CTGCTTTTCT TGCCTGGGTT TGTCTGTCTA TGCTGCCTGG TGCACATCCC 1560  
 TTCTCTTTGC TGCCACTGTT CTATTTTGGG AGTTGTCTTC CGTCTAAGAT GGCTTCTGGG 1620  
 GTTCTATCTT ATTGCACAGA GGTCCCAGAA CAGTGTTTCAT AGGGCACCAT CTGCTCTGCC 1680  
 AAGGGTTTTT TGATGTCTTA CCCTGGGGAT CTTGAGACAG TGGTTACCTT TAGGAGACCC 1740  
 ACCTGGAAC AACCATTAAAG TGACTGCCCA CATTGAGATC AGGGACCATC TTAATAGTAC 1800  
 TCACTGCCAG TCCTCACAAG AGAAGATGAC ACGGGTGCTC TCTTCAGACA CTCCCATACA 1860  
 GGAAGTTGGA AAATGTCTTG GTCACCTGGG TTGTTCCAG GCTACAACTT CTTGGTGTTC 1920  
 CACTAARACC AGRATATCCT AGTTTTTTGG GTTGACTGTT CCCTCCCCAC TTTCTTGAA 1980  
 NCCCAATGCC CNTTTGKTN GGTGCTTCC CTAAAKTT 2019

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 350 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala	Arg	Gly	Gly	Val	Arg	Ala	Glu	Ala	Glu	Asp	Gln	Val	Gly	Met	Ala
1				5					10					15	
Glu	Gly	Gly	Thr	Gly	Pro	Asp	Gly	Arg	Ala	Gly	Pro	Gly	Pro	Ala	Gly
			20					25					30		
Pro	Asn	Leu	Lys	Glu	Trp	Leu	Arg	Glu	Gln	Phe	Cys	Asp	His	Pro	Leu
		35					40					45			

Glu His Cys Asp Asp Thr Arg Leu His Asp Ala Ala Tyr Val Gly Asp  
50 55 60  
Leu Gln Thr Leu Arg Asn Leu Leu Gln Glu Glu Ser Tyr Arg Ser Arg  
65 70 75 80  
Ile Asn Glu Lys Ser Val Trp Cys Cys Gly Trp Leu Pro Cys Thr Pro  
85 90 95  
Leu Arg Ile Ala Ala Thr Ala Gly His Gly Asn Cys Val Asp Phe Leu  
100 105 110  
Ile Arg Lys Gly Ala Glu Val Asp Leu Val Asp Val Lys Gly Gln Thr  
115 120 125  
Ala Leu Tyr Val Ala Val Val Asn Gly His Leu Glu Ser Thr Glu Ile  
130 135 140  
Leu Leu Glu Ala Gly Ala Asp Pro Asn Gly Ser Arg His His Arg Ser  
145 150 155 160  
Thr Pro Val Tyr His Ala Xaa Arg Val Gly Arg Asp Asp Ile Leu Lys  
165 170 175  
Ala Leu Ile Arg Tyr Gly Ala Asp Val Asp Val Asn His His Leu Asn  
180 185 190  
Ser Asp Thr Arg Pro Pro Phe Ser Arg Arg Leu Thr Ser Leu Val Val  
195 200 205  
Cys Pro Leu Tyr Ile Ser Ala Ala Tyr His Asn Leu Gln Cys Phe Arg  
210 215 220  
Leu Leu Leu Gln Ala Gly Ala Asn Pro Asp Phe Asn Cys Asn Gly Pro  
225 230 235 240  
Val Asn Thr Gln Glu Phe Tyr Arg Gly Ser Pro Gly Cys Val Met Asp  
245 250 255  
Ala Val Leu Arg His Gly Cys Glu Ala Ala Phe Val Ser Leu Leu Val  
260 265 270  
Glu Phe Gly Ala Asn Leu Asn Leu Val Lys Trp Glu Ser Leu Gly Pro  
275 280 285  
Glu Ala Arg Gly Arg Arg Lys Met Asp Pro Glu Ala Leu Gln Val Phe  
290 295 300  
Lys Glu Ala Arg Ser Ile Pro Arg Thr Leu Leu Ser Leu Cys Arg Val  
305 310 315 320

Ala Val Arg Arg Ala Leu Gly Lys Tyr Arg Leu His Leu Val Pro Ser  
325 330 335  
Leu Pro Leu Pro Asp Pro Ile Lys Lys Phe Leu Leu Tyr Glu  
340 345 350

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 419 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCATCCATGG CGGAGGGCGG CAGCACGACG GGCGGGCAGG GCCGGGCTCC GCAGGTCGTA 60  
ATCTGAAGGA GTGGCTGAGG GAGCAATTTT GTGATCATCC GCTGGAGCAC TGTGAGGACA 120  
CGAGGCTCCA TGATGCAGCT TACGTCGGGG ACCTCCAGAC CCTCAGGAGC CTATTGCAAG 180  
AGGAGAGCTA CCGGAGCCGC ATCAACGAGA AGTCTGTCTG GTGCTGTGGC TGGCTCCCCT 240  
GCACACCGTT GCGAATCGCG GCCACTGCAG GCCATGGGAG CTGTGTGGAC TTCCTCATCC 300  
GGAAGGGGGC CGAGGTGGAT CTGGTGGACG TAAAAGGACA GACGGCCCTG TATGTGGCTG 360  
TGGTGAACGG GCACCTAGAG AGTACCCAGA TCCTTCTCGA AGCTGGCGCG GACCCCAAC 419

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 595 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAGGAAGAAG AAAAGTGGAC CCTGAGGCCT TGCAGGTCTT TAAAGAGGCC AGAAGTGTTT 60  
CCAGAACCTT GCTGTGTCTG TGCCGTGTGG CTGTGAGAAG AGCTCTTGGC AAAACCGGCT 120  
TCATCTGATT CCTTCGCTGC CTCTGCCAGA CCCATAAAG AAGTTTCTAC TCCATGAGTA 180  
GACTCCAAGT GCTGCGGTTG ATTCCAGTGA GGGAGAAAGT GATCTGCAGG GAGGTGGACA 240

CCGAGCCCTG AGTGCTGTGC TGCTGCTGGT CTCCTGATGG CTGTTGCTGC AGAAGATGTC	300
CTCGTAGACT GTCATTGCTC CTCAGGTGCC TGGGCCGCTG AACAGTCCTT GGGTCATTGT	360
CAGCTGAGAG GCTTATACTA AAGTTATTAT TGT'TTTTCCC AAGTTCTCTG TTCTGGATTT	420
TCAGTTGCAT ATTAATGTAA CGGGCCATGG GGTATGTACA TGTAGGGGCT GAGGTTGGAG	480
GCCTACTAAT TTCCTGTAGG GAAGACTCCC AGCACTTCTG GAACTGTGCT TCTCTTTATT	540
TTTCTACTTC TCAATTTGAT GGTTCGATTA AAGCCTTCTA GTATCTCAAT GAAAA	595

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 896 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 4..396

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTG ATG TCC GCA ATT CTG AAG GTT GGA CAC CAC TGC TGG CTG CCT GTG	48
Met Ser Ala Ile Leu Lys Val Gly His His Cys Trp Leu Pro Val	
1 5 10 15	
ACA TCC GCT GTC AAT CCC CAA AGG ATG CTG AGG CCA CCA CCA ACC GCT	96
Thr Ser Ala Val Asn Pro Gln Arg Met Leu Arg Pro Pro Pro Thr Ala	
20 25 30	
GTT TTC AAC TGT GCC GCT TGC TGC TGT CTG TGG GGG CAG ATG CTG ATG	144
Val Phe Asn Cys Ala Ala Cys Cys Cys Leu Trp Gly Gln Met Leu Met	
35 40 45	
AAT ACA TAC CGT GTA GTT CAG CTT CCT GAG GAG GCC AAG GGC TTG GTG	192
Asn Thr Tyr Arg Val Val Gln Leu Pro Glu Glu Ala Lys Gly Leu Val	
50 55 60	
CCA CCA GAG ATT CTA CAG AAG TAC CAT GGA TTC TAC TCT TCC CTC TTT	240
Pro Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe	
65 70 75	
GCC TTG GTG AGG CAG CCC AGG TCG CTG CAG CAT CTC TGC CGT TGT GCG	288
Ala Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala	
80 85 90 95	



31  
conf

- 147 -

CTC CGC AGT CAC CTG GAG GGC TGT CTG CCC CAT GCA CTA CCG CGC CTT	336
Leu Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu	
100 105 110	
CCC CTG CCA CCG CGC ATG CTC CGC TTT CTG CAG CTG GAC TTT GAG GAT	384
Pro Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp	
115 120 125	
CTG CTC TAC TAGGCTTGCT GCCCTGTGAA CAAAGCAGAC CCCACCCCCA	433
Leu Leu Tyr	
130	
CCCCAAGGGC ATCTCTCAGC AATGAATGAT GCAAGGCGGT CTGTCTTCAA GTCAGGAGTG	493
GACGCCTTGA TCCACACTTG AGAGAAGAGG CCAGATCAGC ACCYGGCTGG TAGTGATNGC	553
AGAGGGCACC TGTGCAGATC TGTGTGCGCA CTGGAAATCT CTAGGCTGAA GGCYAGAGCA	613
AATGGTG CAR GTGTTAGTCC TTGGGANGAG AGACAGANGG TGAGAAAGCA AGACAGAGGT	673
GAGAGTGCAC ATGTCAAGTG GTAGATTGCC TTAAAAGAAA GCTAAAAAAA GAAAAAGATT	733
CGGGCGAACT TCTTTAGGGG TAATGCTGCA GCGTGTAA CTGACTGACC AGCGTCCATA	793
TCTTTGGACC CTTCCCGGGT GAAAAAGCCC CTTATCCTC CAGCGCTCCC CAAGGGTGCT	853
TAGCAATACC GGGTGCTTTT CTGCCGCAA GTGAGTTACC AAA	896

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 130 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met	Ser	Ala	Ile	Leu	Lys	Val	Gly	His	His	Cys	Trp	Leu	Pro	Val	Thr
1				5				10						15	
Ser	Ala	Val	Asn	Pro	Gln	Arg	Met	Leu	Arg	Pro	Pro	Pro	Thr	Ala	Val
			20					25					30		
Phe	Asn	Cys	Ala	Ala	Cys	Cys	Cys	Leu	Trp	Gly	Gln	Met	Leu	Met	Asn
		35					40					45			
Thr	Tyr	Arg	Val	Val	Gln	Leu	Pro	Glu	Glu	Ala	Lys	Gly	Leu	Val	Pro
	50					55					60				

Pro Glu Ile Leu Gln Lys Tyr His Gly Phe Tyr Ser Ser Leu Phe Ala  
 65 70 75 80  
 Leu Val Arg Gln Pro Arg Ser Leu Gln His Leu Cys Arg Cys Ala Leu  
 85 90 95  
 Arg Ser His Leu Glu Gly Cys Leu Pro His Ala Leu Pro Arg Leu Pro  
 100 105 110  
 Leu Pro Pro Arg Met Leu Arg Phe Leu Gln Leu Asp Phe Glu Asp Leu  
 115 120 125  
 Leu Tyr  
 130

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 436 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GTGGGGGCGT CATCATGACC TCCTCTAGGG CTCTGCAACA TGACTCCTGT GGTGCAAATC	60
AACAAATTGT TCACTGATGA ATCCACAAGG ATCTCTGGGC CTACAACCAG GTCCTGGTCC	120
ACATGACTGT CGTCTTCGGA GAAGGCACCA CTCGCCCCCG GCAGGTACGG CTGACACCTC	180
CATGGGAGAA GACGTATCCA GGCAGCAGCT GCGCGGCCCT TCAAGAGGGC ACATCCCGTC	240
ATCTAAAGGC ACGGTGTACT GAAGGTAGTC CTGAGACATG AGTCCGATTA CTACAGGCAC	300
GTGTTCTCTC AGGTGGAGGC TCAGGTCCCC GGGTGAGCTG GGGCTGCAGC GGGACTCAGG	360
GCGCGGCTCT GGCTGCAGGT CTCGCAGCTC CCTGGGCTGT AGCTCCCGCA GATCCTTGCG	420
CACACCGTTG ACTGGT	436

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2180 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TTAATAGTAC CTACATAGTA GAAAATTATA ACTCCACTTT AAAACAATGT TTTCTTTCTA	60
TTCAAATCAA TTTAAACTT TTTATAAACA TTAATGTTGC AAGAGAATCC AGTCCATTTA	120
TGAAAATTAG TTGACAATCA AGTTCACCCA AGAAAATGTT GACTAAGCTA AAGAAATCAC	180
AGATAAAACA TTTTACCAAA AGGATAGGTA ACACACAAAA AAATGCTATC ACAGGAAGCT	240
ATGATCATCT AATATTTCTT TAATAATAAT TCTAGTTCCA TAGGTTTTCA TGTTATGCCA	300
ATTTGTACCC GAGTTTAATT ACAGAAAAGG CAACAATTC TAAATTGGTG GTATACATTT	360
CTTTACAATT TTTTAATGTA AGGCCATTTA TTAAAATAGA CAACTAGAA GATGAAAACG	420
AAGGCAACAG AAAAATTCAA CTTTTCACAA CCAAAGAAT TAGCACAACC TTAGAAATAA	480
TTTAGAAAAA AGTGTTGTTA AAAGATATGT TGCAGATCTC CGTTCCATTA CCCAAGATTA	540
TGTCAATTCA CGATTCTAAA TAAATCTTTT TAAAGTAAGA GATTAAAAAC TCATCTTCAG	600
TGTATATGTA AATTCCGTGG TTTTATCACA CAGGTATGTT TATTCAACAC TGCTTTGGAA	660
ATGGACCATT TAAAAGGACA TGGCAATTC CATTCTGTTA AGTTTCATTC AACCTTTACT	720
TAGGGGTGTA TTACCACATG AAATGTGCTT TTAATGCATA AAAATCACAG TGGATTAGCC	780
AGCAAAAGGG ACTGGGCGGG GGGGCGATTG AGGAGAATTT GATAATTCAC ATTGTGATTA	840
TTCTGCACAT TGATGAAACA TAATTCACAC CTCTAAAACC TCAAGACTTC CCTTTTTTAA	900
AGAACCAAAA TAAACCCAAG ACACCTTGCT GACACTTCCC CACCCTAAA CAACTGATG	960
ACTCTTTTAC ACATAAACT GAAATAGTTA TGGCAGCAAA AGATTTTGAT GGCAATGAAA	1020
GTTTGTAAC TGTATTTCAA TCTCTTGTTT TTATTCCCAA AGTGCAAGAT GCAGGGTTCT	1080
CAATCTTTCA GTAGTGCTTC TCCTGTAAAT AATCCTTCAT TTTGTTTGGC AAAGGCAGTT	1140
TCTGAATTAA GTCTATTCTG GTATACTGAC GTATAACAAA ACGACACAGG TACTGCAACG	1200
AGCGCACCTA TGAACCCCGG AACACTGGTT GGCAAGTTCT GACGGAAGTG CAGATTCCAG	1260
GCAGCGAGAC CTTGAATAAC AAAAAGCTCC CATTTTCAGA GTCCCTGATT GAATGCTCCA	1320
ATTAGATCAA CTATGGACGT ATGTCCTTCC ACATCGGCTG TTCATAAAAG CTAAACCTAC	1380
CATTTGAGTG CTCAATTCTA GTGTGAAGTG TTTTACCATG GGAGCGAAAG TCACAGCTTA	1440
AAAGGTAACG GTCGTCAGAA CTGTCCCGAA CAAGAAAAGA ACCATCTGGC ACGTTTGCTA	1500

B1  
conf.

- 150 -

GCTTCCCTTC	TGCCTCCCAA	CGTGTGATTG	GTCCCCAGTA	CCATCCTTGC	TTTGCAAGTT	1560
TTTTCAGCTC	CTCTGTAAGG	CTTGTCACAA	CCATGGGACC	ACTACTTTGC	ACTGAGTCAT	1620
AAACTCTTGC	AACCCCAGGA	GCAGAGTTCG	GATCAAAATT	CAAATGACAG	CGCATAACTT	1680
TCAGCCACGT	GGGGCTTTCT	GTCCAGTGAG	TCCACTGAAA	GTTCCCTTTT	GGGATTTGGA	1740
TTATTCTGTC	ATTGGAGTAA	CCAATGGTGA	AGATTGGAGG	GACATCCATC	GTGAACCCGC	1800
TCTCCGGGGT	TCTGCAACAT	GACTCCCGTG	GTGCCAATCA	ACAAGCCATT	CACCGGACTG	1860
ATCCACGAAG	ATCTCTGGGG	CGACAAC TAG	GTCCTGGTCT	ACCTGACTCT	CATCCTCGGG	1920
GAAAGCGCGC	CCTCCCCTT	GAGGAGGAAC	CGCAGAGACT	TCCATGGGAG	AAGAGCTGTC	1980
CAGACAATAG	CTCCGTGATC	CTTCCAAAGG	ATACATCCCC	TCATCTAAAG	GCACAGTATA	2040
CTGAATGTAG	TCCTGAGGCA	TAAGTCCAAT	AACGACAGGC	ACATGTTTAT	CCAGGTGAAG	2100
ATGCAGGTCT	CCATTATGAG	AAGCCGAGCT	CTTCAGTGAA	TTGGCTTGCT	CCTGGCACGT	2160
GGTCTCAGAC	TGGAGGTCGT					2180

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2649 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GGCACGAGGC	TGTGTCCAGC	ACACAGAGAG	GGCCCGGCCA	TCTGCTTTGG	TTCAGAGCCC	60
TGTGTCTGTC	TGTCACCTAG	ACTCTTCCTC	CCGGCTCGCA	GCTCACCTC	CATCCTCCTT	120
ACTGGCTCCA	GCATGACTCG	CTTCTCTTAT	GCAGAGTACT	TTGCTCTGTT	TCACTCTGGC	180
TCTGCACCTT	CCAGGTCCCC	TTTGTCTCCC	GAGAACCAC	CGGCCCGCGC	ACCCCTGGGT	240
CTGTTCCAAG	GGGTCATGCA	GAAGTATAGC	AGCAACCTGT	TCAAGACCTC	CCAGATGGCG	300
GCTATGGACC	CCGTGCTGAA	GGCCATCAAG	GAAGGGGATG	AAGAGGCCTT	GAAGATCATG	360
ATCCAGGATG	GGAAGAATCT	TGCAGAGCCC	AACAAGGAGG	GCTGGCTGCC	GCTCCACGAG	420
GCTGCCTACT	ATGGCCAGCT	GGGCTGCCTG	AAAGTCCTGC	AGCAAGCCTA	CCCAGGGACC	480

ATTGACCAAC	GCACACTGCA	GGAAGAGACA	GCATTATACC	TGGCCACATG	CAGAGAACAC	540
CTGGATTGCC	TCCTGTCGCT	GCTCCAGGCG	GGGGCAGAGC	CTGACATCTC	TAACAAATCC	600
AGGGAGACTC	CACTTTACAA	AGCCTGTGAG	CGCAAGAACG	CGGAGGOGGT	GAGGATATTG	660
GTGCGATACA	ACGCAGACGC	CAACCACCGC	TGTAACAGGG	GCTGGACCGC	ACTGCACGAG	720
TCTGTCTCCC	GCAATGACCT	GGAGGTCATG	GAGATCCTAG	TGAGTGGCGG	GGCCAAGGTG	780
GAGGCCAAGA	ATGTCTACAG	CATCACCCCT	TTGTTTGTGG	CTGCCCAGAG	TGGGCAGCTG	840
GAGGCCCTGA	GGTTCCTGGC	CAAGCATGGT	GCAGACATCA	ACACGCAGGC	CAGTGACAGT	900
GCATCAGCCC	TCTACGAGGC	CAGCAAGAAT	GAGCATGAAG	ACGTGGTAGA	GTTTCTTCTC	960
TCTCAGGGCG	CCGATGCTAA	CAAAGCCAAC	AAGGACGGCC	TGCTCCCCCT	GCATGTTGCC	1020
TCCAAGAAGG	GCAACTATAG	AATAGTGCAG	ATGCTGCTGC	CTGTGACCAG	CCGCACGCGC	1080
GTGCGCCGTA	GCGGCATCAG	CCCGCTGCAC	CTAGCGGCCG	AGCGCAACCA	CGACGCGGTG	1140
CTGGAGGCGC	TGCTGGCCGC	GCGCTTCGAC	GTGAACGCAC	CTCTGGCTCC	CGAGCGCGCC	1200
CGCCTCTACG	AGGACCGCCG	CAGTTCTGCG	CTCTACTTCG	CTGTGGTCAA	CAACAATGTG	1260
TACGCCACCG	AGCTGTTGCT	GCTGGCGGGG	GCGGACCCCA	ACCGCGATGT	CATCAGCCCT	1320
CTGCTCGTGG	CCATCCGCCA	CGGCTGCCTG	CGCACCATGC	AGCTGCTGTT	GGACCATGGC	1380
GCCAACATCG	ACGCCTACAT	CGCCACTCAC	CCCACCGCCT	TTCCAGCCAC	CATCATGTTT	1440
GCCATGAAGT	GCCTGTCGTT	ACTCAAGTTC	CTTATGGACC	TCGGCTGCGA	TGGCGAGCCC	1500
TGCTTCTCCT	GCCTGTACGG	CAACGGGCCG	CACCACCCGC	CCCGCGACCT	GGCCGCTTCC	1560
ACGACGCACC	CGTGGACGAC	AAGGCACCTA	GCGTGGTGCA	GTTCTGTGAG	TTCTGTGCGG	1620
CCCCGGAAGT	GAGCCGCTGG	GCGGGACCCA	TCATCGATGT	CCTCCTGGAC	TATGTGGGCA	1680
ACGTGCAGCT	GTGCTCCCGG	CTGAAGGAGC	ACATCGACAG	CTTTGAGGAC	TGGGCTGTCA	1740
TCAAGGAGAA	GGCAGAACCT	CCGAGACCTC	TGGCTCACCT	CTGCCGGCTG	CGGGTTTCGA	1800
AGGCCATAGG	AAAATACCGG	ATAAACTCC	TGGACACACT	GCCGCTTCCC	GGCAGGCTAA	1860
TCAGATACTT	GAAATATGAG	AATACACAGT	AACCAGCCTG	GAGAGGAGAT	GTGGCCTTCA	1920
GACTGTTTCC	GGGACGCCCC	AGGTGGCCTG	CATCCAGGAC	CCCCTGGGGT	CAGAACAGGT	1980
GTGACCTTGC	TGGTTCTTTG	CTGGAGCTTC	ACCCAAAGTG	AGAACCTGAT	GTGGGGAGTG	2040

GACGTGGAAC CTCTGCTTTC ACACTGTCAG CGGATCGCAG ACCCGCTCTG CTTCTGGCCA	2100
TAGCCAGAGA CCTTCAACCT GGGGCCAGGG GAGAGCTGGT CTGGGCAAGG TGGCCCAGGC	2160
AGGAATCCTG GCCTTAAGCT GGAGAACTTG TAGGAATCCC TCACTGGACC CTCAGCTTTC	2220
AGGCTGCGAG GGAGACGCCC AGCCCAAGTA TTTTATTTCC GTGACACAAT AACGTTGTAT	2280
CAGAAAAAAA AAAAACATG GGCGCAGCTT ATTCCTTAGT AGGGTATTTA CTTGCATGCG	2340
CGCTTAAAGC TACTGGAAAC ATGCGTTCCA CTATGCTTGA GAATCCCCTT GCACTGGTAA	2400
ACGAGAGCCG ACGTGCTTCA AGGTTGGATT TTTGGTTGCC CTTTGGCGT TCCGCGGGTT	2460
TGTCCGACGT AATTGACCCC GTGTTTTGTC ACTTTCGAGT GTTCCGACTA TTGGGGGGCT	2520
TTTGGTTGTC CCCAAAATTG TGGGTGGTGT GCGGACGCCA CGAGAAGTGG TTCATGGGCG	2580
ATAATCATTA CTGGAGAATG TAGAGCGGCG GTTTTACGAA TAAATATTTT TTAAGCCGCC	2640
TTCCCAAAA	2649

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 495 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCTCCTGAGA GTTCGCCGGC CCGGGCCCCA TGGGTGTTC CAAGGGGTCA TGCAGAAATA	60
CAGCAGCAGC TTGTTCAAGA CCTCCCAGCT GGCGCCTGCG GACCCCTTGA TAAAGGCCAT	120
CAAGGATGCG ATGAAGAGGC CTTGAAGACC ATGATCAAGG AAGGGAAGAA TCTCGCAGAG	180
CCCAACAAGG AGGGCTGGCT GCCGCTGCAC GAGGCCGCAT ACTATGGCCA GGTGGGCTGC	240
CTGAAAGTCC TGCAGCGAGC GTACCCAGGG ACCATCGACC AGCGCACCTT GCAGGAGGAA	300
ACAGCCGTTT ACTTGGCAAC GTGCAGGGGC CACCTGGACT GTCTCCTGTC ACTGCTCCAA	360
GCAGGGGCAG AGCGGGACAT CTCCAACAAA TCCCGAGAGA ACCGCTCTAC AAAGCCTGTG	420
AGCGCAAGAA CGCGGAAGCC GTGAAGATTC TTGGTGCAGC ACAACGCAGA CACCAACAAC	480
GCTGCAACCG GGCTG	495

(2) INFORMATION FOR SEQ ID NO:34:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 709 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GTGCAGCTCT GCTCGCGGCT GAAGGAACAC ATCGACAGCT TTGAGGACTG GGCCGTCATC	60
AAGGAGAAGG CAGAACCTCC AAGACCTCTG GCTCACCTTT GCCGACTGCG GGTTCGAAAG	120
GCCATTGGGA AATACCGTAT AAAACTCCTA GACACCTTGC CGCTCCCAGG CAGGCTGATT	180
AGATACCTGA AATACGAGAA CACCCAGTAA CTGGGGCCAC GGGGAGAGAG GAGTAGCCCC	240
TCAGACTCTT CTTACTAAGT CTCAGGACGT CGGTGTTCC AACTCCAAGG GGACCTGGTG	300
ACAGACGAGG CTGCAGGCTG CCTCCCTCTC AGCCTGGACA GCTACCAGGA TCTCACTGGG	360
TCTCAGGGCC CAGAGCTTTG GCCAGAGCAG AGAACAGAAT GTGTCAAGGA GAAGAATCAT	420
TTGTTTACAA ACTGATGAGC AGATCCCAGA CCTTCTCTAC CTTCAGGAAT GGCAGAAACC	480
TCTATTCCTG GGGCCAGGGC AGAGCTTGAG GTGTTCTGGG GAAGGTGGTG CTCAGAGCCT	540
TCCCTGTGCC CCTCCACTTG TTCTGGAAAA CTCACCACTT GACTTCAGAG CTTTCTCTCC	600
AAAGACTAAG ATGAAGACGT GGCCCAAGGT AGGGGGTAGG GGGAGCCTGG GTCTTGAGG	660
GCTTTGTAA GTATTAATAT AATAAATGTT ACACATGTGA AAAAAAAAAA	709

(2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 848 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1..624

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTG	GAG	AAG	TGT	GGT	TGG	TAT	TGG	GGG	CCA	ATG	AAT	TGG	GAA	GAT	GCA	48
Leu	Glu	Lys	Cys	Gly	Trp	Tyr	Trp	Gly	Pro	Met	Asn	Trp	Glu	Asp	Ala	
1				5				10						15		
GAG	ATG	AAG	CTG	AAA	GGG	AAA	CCA	GAT	GGT	TCT	TTC	CTG	GTA	CGA	GAC	96
Glu	Met	Lys	Leu	Lys	Gly	Lys	Pro	Asp	Gly	Ser	Phe	Leu	Val	Arg	Asp	
			20					25					30			
AGT	TCT	GAT	CCT	CGT	TAC	ATC	CTG	AGC	CTC	AGT	TTC	CGA	TCA	CAG	GGT	144
Ser	Ser	Asp	Pro	Arg	Tyr	Ile	Leu	Ser	Leu	Ser	Phe	Arg	Ser	Gln	Gly	
		35					40					45				
ATC	ACC	CAC	CAC	ACT	AGA	ATG	GAG	CAC	TAC	AGA	GGA	ACC	TTC	AGC	CTG	192
Ile	Thr	His	His	Thr	Arg	Met	Glu	His	Tyr	Arg	Gly	Thr	Phe	Ser	Leu	
	50					55					60					
TGG	TGT	CAT	CCC	AAG	TTT	GAG	GAC	CGC	TGT	CAA	TCT	GTT	GTA	GAG	TTT	240
Trp	Cys	His	Pro	Lys	Phe	Glu	Asp	Arg	Cys	Gln	Ser	Val	Val	Glu	Phe	
65					70					75					80	
ATT	AAG	AGA	GCC	ATT	ATG	CAC	TCC	AAG	AAT	GGA	AAG	TTT	CTC	TAT	TTC	288
Ile	Lys	Arg	Ala	Ile	Met	His	Ser	Lys	Asn	Gly	Lys	Phe	Leu	Tyr	Phe	
				85					90					95		
TTA	AGA	TCC	AGG	GTT	CCA	GGA	CTG	CCA	CCA	ACT	CCT	GTC	CAG	CTG	CTC	336
Leu	Arg	Ser	Arg	Val	Pro	Gly	Leu	Pro	Pro	Thr	Pro	Val	Gln	Leu	Leu	
			100					105					110			
TAT	CCA	GTG	TCC	CGA	TTC	AGC	AAT	GTC	AAA	TCC	CTC	CAG	CAC	CTT	TGC	384
Tyr	Pro	Val	Ser	Arg	Phe	Ser	Asn	Val	Lys	Ser	Leu	Gln	His	Leu	Cys	
		115					120					125				
AGA	TTC	CGG	ATA	CGA	CAG	CTC	GTC	AGG	ATA	GAT	CAC	ATC	CCA	GAT	CTC	432
Arg	Phe	Arg	Ile	Arg	Gln	Leu	Val	Arg	Ile	Asp	His	Ile	Pro	Asp	Leu	
	130					135					140					
CCA	CTG	CCT	AAA	CCT	CTG	ATC	TCT	TAT	ATC	CGA	AAG	TTC	TAC	TAC	TAT	480
Pro	Leu	Pro	Lys	Pro	Leu	Ile	Ser	Tyr	Ile	Arg	Lys	Phe	Tyr	Tyr	Tyr	
145					150					155					160	
GAT	CCT	CAG	GAA	GAG	GTA	TAC	CTG	TCT	CTA	AAG	GAA	GCG	CAG	CGT	CAG	528
Asp	Pro	Gln	Glu	Glu	Val	Tyr	Leu	Ser	Leu	Lys	Glu	Ala	Gln	Arg	Gln	
				165					170				175			
TTT	CCA	AAC	AGA	AGC	AAG	AGG	TGG	AAC	CCT	CCA	CGT	AGC	GAG	GGG	CTC	576
Phe	Pro	Asn	Arg	Ser	Lys	Arg	Trp	Asn	Pro	Pro	Arg	Ser	Glu	Gly	Leu	
			180					185					190			
CCT	GCT	GGT	CAC	CAC	CAA	GGG	CAT	TTG	GTT	GCC	AAG	CTC	CAG	CTT	TGA	624
Pro	Ala	Gly	His	His	Gln	Gly	His	Leu	Val	Ala	Lys	Leu	Gln	Leu		
		195					200					205				



AGAACCAAAT TAAGCTACCA TGAAAAGAAG AGGAAAAGTG AGGGAACAGG AAGGTTGGGA	684
TTCTCTGTGC AGAGACTTTG GTTCCCCACG CAAGCCCTGG GGCTTGGAAG AAGCACATGA	744
CCGTA CTCTG CGTGGGGCTC CACCTCACAC CCACCCCTGG GCATCTTAGG ACTGGAGGGG	804
CTCCTTGGA AACTGGAAGA AGTCTCAACA CTGTTTCTTT TTCA	848

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 207 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Leu	Glu	Lys	Cys	Gly	Trp	Tyr	Trp	Gly	Pro	Met	Asn	Trp	Glu	Asp	Ala	
1				5					10					15		
Glu	Met	Lys	Leu	Lys	Gly	Lys	Pro	Asp	Gly	Ser	Phe	Leu	Val	Arg	Asp	
			20					25					30			
Ser	Ser	Asp	Pro	Arg	Tyr	Ile	Leu	Ser	Leu	Ser	Phe	Arg	Ser	Gln	Gly	
		35					40					45				
Ile	Thr	His	His	Thr	Arg	Met	Glu	His	Tyr	Arg	Gly	Thr	Phe	Ser	Leu	
	50					55					60					
Trp	Cys	His	Pro	Lys	Phe	Glu	Asp	Arg	Cys	Gln	Ser	Val	Val	Glu	Phe	
65					70				75						80	
Ile	Lys	Arg	Ala	Ile	Met	His	Ser	Lys	Asn	Gly	Lys	Phe	Leu	Tyr	Phe	
				85					90					95		
Leu	Arg	Ser	Arg	Val	Pro	Gly	Leu	Pro	Pro	Thr	Pro	Val	Gln	Leu	Leu	
			100				105						110			
Tyr	Pro	Val	Ser	Arg	Phe	Ser	Asn	Val	Lys	Ser	Leu	Gln	His	Leu	Cys	
	115						120					125				
Arg	Phe	Arg	Ile	Arg	Gln	Leu	Val	Arg	Ile	Asp	His	Ile	Pro	Asp	Leu	
	130				135					140						
Pro	Leu	Pro	Lys	Pro	Leu	Ile	Ser	Tyr	Ile	Arg	Lys	Phe	Tyr	Tyr	Tyr	
145					150				155					160		
Asp	Pro	Gln	Glu	Glu	Val	Tyr	Leu	Ser	Leu	Lys	Glu	Ala	Gln	Arg	Gln	
			165					170						175		

B1  
cont

- 156 -

Phe Pro Asn Arg Ser Lys Arg Trp Asn Pro Pro Arg Ser Glu Gly Leu  
180 185 190  
Pro Ala Gly His His Gln Gly His Leu Val Ala Lys Leu Gln Leu  
195 200 205

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GTTCCAAGCC TAACCCATCT TTGTCGTTTG GAAATTCGGG CCAGTCTAAA AGCAGAGCAC	60
CTTCACTCTG ACATTTTTCAT CCATCAGTTG CCACTTCCCA GAAGTCTGCA GAACTATTTG	120
CTCTATGAAG AGGTTTTAAG AATGAATGAG ATTCTAGAAC CAGCAGCTAA TCAGGATGGA	180
GAAACCAGCA AGGCCACCTG ACACAGGTCC TTTAATTCTG TTTAGTCACA AAAGACGGCT	240
TGTGTGACTG TTTGGATTTG GTGATCAAAT GTCCATGTTT ACAGTTGCTT TTCCCAGTTT	300
GTGTCTTTCC CAATATTGTG AACCTTATCC ATCTTGCCTT ACTCAGTTTT ATTTCTAGTG	360
CACTTTGTTG TGTATTATTT GTTTACCTGA CCATTTTCTA CTTTATTCTG CTAATAAACT	420
GTAATTCTGA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAA	464

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 747 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGGGATCGAA AGCGGGGGCT TCTGGGACGC AGCTCTGGAG ACGCGGCCTC GGACCAGCCA	60
TTTCGGTGTA GAAGTGGCAG CACGGCAGAC TGGTCAAACA AATGGATTTT ACAGAGGCTT	120
ACGCGGACAC GTGCTCTACA GTTGGACTTG CTGCCAGGGA AGGCAATGTT AAAGTCTTAA	180

GGAAACTGCT	CAAAAAGGGC	CGAAGTGTCTG	ATGTTGCTGA	TAACAGGGGA	TGGATGCCAA	240
TTCATGAAGC	AGCTTATCAC	AACTCTGTAG	AATGTTTGCA	AATGTTAATT	AATGCAGATT	300
CATCTGAAAA	CTACATTAAG	ATGAAGACCT	TTGAAGGTTT	CTGTGCTTTG	CATCTCGCTG	360
CAAGTCAAGG	ACATTGGAAA	ATCGTACAGA	TTCTTTTAGA	AGCTGGGGCA	GATCCTAATG	420
CAACTACTTT	AGAAGAAACG	ACACCATTGT	TTTTAGCTGT	TGAAAATGGA	CAGATAGATG	480
TGTTAAGGCT	GTTGCTTCAA	CACGGAGCAA	ATGTTAATGG	ATCCCATTCT	ATGTGTGGAT	540
GGAAGTCCTT	GCACCAGGCT	TCTTTTCAGG	AAAATGCTGA	GATCATAAAA	TTGCTTCTTA	600
GAAAAGGAGC	AAACAAGGAA	TGCCAGGATG	ACTTTGGAAT	CACACCTTTA	TTTGTGGCTG	660
CTCAGTATGG	CCAAGCTAGA	AAGCTTTGAA	GCATACTTAT	TTCATCCGGG	TGCAAATGTC	720
AATTGTCAAG	CCTTGGACAA	AGCTACC				747

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1018 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CACAAATGGG	ACCATACAAA	AATCTTGGAC	TTGTTAATAA	CCACTTACTA	ACCGGGACCT	60
GTGACACTGG	GCTAAACAAA	GTAAGTCCCT	GTTTACTCAG	CAGTGTTTGG	GGGACATGAA	120
GGATTGCCTA	GAAATATTAC	TCCGGAATGG	TCTACAGCCC	AGACGCCAG	GCGTGCCTTG	180
TTTTTGGATT	CAGTTCTCCT	GTGTGCATGG	CTTTCCAAAA	GGAGGTGGAG	CTGTAGTTCT	240
TTGGAATTGT	GAACATTCTT	TTGAAATATG	GAGCCCAGAT	AAATGAACTT	CATTTGGCAT	300
ACTGCCTGAA	GTACGAGAAG	TTTTCGATAT	TTCGCTACTT	TTTGAGGAAA	GGTTGCTCAT	360
TGGGACCATG	GAACCATATA	TATGAATTTG	TAAATCATGC	AATTAAAGCA	CAAGCAAAAT	420
ATAAGGAGTG	GTTGCCACAT	CTTCTGGTTG	CTGGATTTGA	CCCACTGATT	CTACTGTGCA	480
ATTCTTGGAT	TGACTCAGTC	AGCATTGACA	CCCTTATCTT	CACTTTGGAG	TTTACTAATT	540
GGAAGACACT	TGCACCAGCT	GTTGAAAGGA	TGCTCTCTGC	TCGTGCCTCA	AACGCTTGGA	600

TTCTACAGCA	ACATATTGCC	CACTGTTCCA	TCCCTGACCC	ATCTTTGTCG	TTTGGAAATT	660
CGGTCCAGTC	TAAAATCAGA	ACGTCTACGG	TCTGACAGTT	ATATTAGTCA	GCTGCCACTT	720
CCCAGAAGCC	TACATAATTA	TTTGCTCTAT	GAAGACGTTT	TGAGGATGTA	TGAAGTTCCA	780
GAACTGGCAG	CTATTCAAGA	TGGATAAATC	AGTGAAACTA	OTTAACACAG	CTAATTTTTT	840
TCTCTGAAAA	ATCATCGAGA	CAAAAGAGCC	ACAGAGTACA	AGTTTTTATG	ATTTTATAGT	900
CAAAAGATGA	TTATTGATTG	TCAGATAGGT	TAGGTTTTGG	GGGGCCAGTA	GTTTCAGTGAG	960
AATGTTTATG	TTTACAACTA	GCCTTCCCAG	TAAAAAATAA	AAAAAAAAAA	AAAAAAAAAA	1018

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1897 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CGGGGGGCTG	GGACCTGGGG	CGTAACCGTC	TCTACCACGA	CGGCAAGAAC	CAGCCAAGTA	60
AAACATACCC	AGCCTTTCTG	GAGCCGGACG	AGACATTCAT	TGTCCCTGAC	TCCTTTTTCG	120
TGGCCCTGGA	CATGRATGAT	GGGACCTTAA	GTTTCATCGT	GGATGGACAG	TACATGGGAG	180
TGGCTTTCCG	GGGACTCAAG	GGTAAAAAGC	TGTATCCTGT	AGTGAGTGCC	GTCTGGGGCC	240
ACTGTGAGAT	CCGCATGCGC	TACTTGAACG	GACTTGATCC	TGAGCCCCTG	CCACTCATGG	300
ACCTGTGCCG	GCGTTTCGGT	CGCCTAGCGC	TGGGAAAAGA	GCGCCTGGGT	GCCATCCCCG	360
CTCTGCCGCT	ACCTGCCTCC	CTCAAAGCCT	ACCTCCTCTA	CCAGTGATCC	ACATCCCAGG	420
ACCGCCATAC	GACAGCCATC	TGGTGCCAAR	TCACTGAGCC	CGTTGGGGTC	CGCCGACCCC	480
TGCGCCTGGG	ATGGAAGCCC	ACCTCAGCCA	TGGGCAGACG	TGCCCCCTCA	TCCTACCGGC	540
TGCCTCTGCT	GGGGGAACCT	ATGCCAACGG	ACTTCTCCCT	TCCCAACACT	GGCTGAAGCA	600
GCAGCACCCA	GGCCCTTCCC	TGAACCAGAT	GCAGAGAATA	AACTATGAAA	ACCTCTCTCA	660
GGCGCCTTCT	GCTCTCAGGT	GGAGTGGGCT	GCCCCCACT	CTCTGCAGAG	AGAGGCTACA	720
CCCACCTGGG	GGGTCCTGGG	AGGTAAGACT	AGTAGGAGGT	GCCAGGGCTG	ARTCCAAAAG	780

CONF.

- 159 -

CAGGAATGGC	CAGGAMCAGG	CCATACAGAT	GAAGCTCAGG	ATGTCACATA	CCATGGACAM	840
TGAGACAGAA	CCCCAGGTTG	GAMTTCCCTT	GGGCCAACGA	GTGCCAGCTT	TAATGTCAGC	900
TGCMGGTGCT	CTGTGGCCTG	TATTTATTCT	TTAAACAGTA	GCAAAGGCCA	TTTATTTATT	960
CCACTTAGAA	AGGAAACCTT	GGTGGGTGGY	TTCCCTCGAT	GTGCTTTCCC	CCACCTCCCT	1020
GGAATGTGTG	TGCCACACCT	GTCCTTGTCC	CAGGCCAGGA	CTGTGGCACA	TGAGCTGGTG	1080
TGCACAGATA	CACGTATGTC	GTCGTGCATG	ACCCCTGACT	AGTTCCTAAG	TAGCCCTGCA	1140
CCAAGCACCA	GAGCAGACCC	CAAGAGAGGC	CCGTGCAAGT	CCCCATGTCC	CCAGGTCCCT	1200
GCTTCTGTTG	CCTTGGGACT	CATACACCGG	CACACGTGTT	TCAGCCTCTT	GACTTCCATG	1260
AGCTTCGAAT	TTTGCCCCCG	ATTCTTCTGA	TATTTCCCAT	TGGCATCCTC	CAAAGCTCTG	1320
GGCCTGGAGG	GCATTAGGAC	ACATGGAATG	AGTGGGGTCT	CCAGCCCCTG	GGAAAGCCAC	1380
TGGCAAGGCA	GGATTAGAAA	GACCAAGAGC	AGGGTGGGGC	GCCATGAAGC	CTGTATGCCT	1440
CTCAGGCTCA	AGACCCCGCC	ACACACCCAC	TCAAGCCTCA	GAAGTGGTGT	GTAGGGCAGC	1500
CCCAGGAGAG	GAATGCCTGT	CCTAGCAGCA	CGTACATGGA	GCACCCACAC	TGTGCTCCAG	1560
CCCTCTGGCT	GTTTCTCTTG	CTCTAGAATC	AACTCCCTAC	ATTGGGAATG	TAGCCATTTG	1620
GTAGAGGACT	TGCCTAGCCT	GCAGGAAGCT	CACGTTCCAT	CCCCTGCACC	AAGGAGAATC	1680
AAAGCTCAGG	AGGCTGAGGC	AGGAGGATTG	CTGTCAGTGG	TGTACAGAGG	TCATGGCCAT	1740
CCTGGGCTAT	ATTAAACCTT	GTCCTTTAAG	AAAAAGAAAA	GAAATCAACT	TCCATTGAAT	1800
CTGAGTTCTG	CTCATTTCTG	CACAGGTACA	ATAGATGACT	TKATTTGTTG	AAAAATGKTT	1860
AATATATTTA	CMTATATATA	TATTTGTAAG	AAGCATT			1897

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 134 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Gly	Gly	Trp	Asp	Leu	Gly	Arg	Asn	Arg	Leu	Tyr	His	Asp	Gly	Lys	Asn	Gln
1				5				10					15			

- 160 -

Pro Ser Lys Thr Tyr Pro Ala Phe Leu Glu Pro Asp Glu Thr Phe  
20 25 30  
Ile Val Pro Asp Ser Phe Phe Val Ala Leu Asp Met Xaa Asp Gly Thr  
35 40 45  
Leu Ser Phe Ile Val Asp Gly Gln Tyr Met Gly Val Ala Phe Arg Gly  
50 55 60  
Leu Lys Gly Lys Lys Leu Tyr Pro Val Val Ser Ala Val Trp Gly His  
65 70 75 80  
Cys Glu Ile Arg Met Arg Tyr Leu Asn Gly Leu Asp Pro Glu Pro Leu  
85 90 95  
Pro Leu Met Asp Leu Cys Arg Arg Ser Val Arg Leu Ala Leu Gly Lys  
100 105 110  
Glu Arg Leu Gly Ala Ile Pro Ala Leu Pro Leu Pro Ala Ser Leu Lys  
115 120 125  
Ala Tyr Leu Leu Tyr Gln  
130

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 265 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAGGGTAAAA AACTGTATCC TGTAGTGAGT GCCGTCTGGG GCCACTGTAG ATCCGAATGC 60  
GCTACTTGAA CGGACTCGAT CCCGAGACTG CCGCTCATGG ATTTGTGCCG TCGCTCGGTG 120  
CGCCTGGCCC TGGGGAGGGA GCGCCTGGGG GAGAACCACA CCTGCCGCTG CCGGCTTCCC 180  
TCAAGGCCTA CCTCCTCTAC CAGTGACGTT CGCCATCATA CCGCCAGCGC GACAGCCACC 240  
TGGTGCCAAC TCACTGAGCC GCCTG 265

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2438 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AAGTGGCGGC	GGTCCCTGGA	GAGCAGGCGG	AGGCAGCGGC	AAGTCTGACT	CTGGGCTGAC	60
CGTGGAGCCG	GGGCGGGGGC	TGACAGCCAG	GCCTCCGCCT	GGCGGGAGCC	GCACGAGGAG	120
CGGGAGTGGC	CGGGCCTCTC	TTCCGCGCTT	GAGCGAGCGC	CGGGTGATGG	CGGTGGTGAT	180
GGCGGCAGGC	GCTCGGACAG	CTCCGCTTGA	GCTGAGCTCG	GAGAGATCCG	TCCAGAAAGT	240
GCCCAGAAGA	AAC TTCCTCT	TAGAAAAGCT	GAAAAACACA	RTATTTATAA	CACTGGAAAT	300
TGTAAAGAAT	TTGTTTAAAA	TGGCTGAAAA	CAATAGTAAA	AATGTAGATG	TACGGCCTAA	360
AACAAGTCGG	AGTCGAAGTG	CTGACAGGAA	GGATGGTTAT	GTGTGGAGTG	GAAAGAAGTT	420
GTCTTGGTCC	AAAAAGAGTG	AGAGTTGTTC	TGAATCTGAA	GCCATAGGTA	CTGTTGAGAA	480
TGTTGAAATT	CCTCTAAGAA	GCCAAGAAAG	GCAGCTTAGC	TGTTCGTCCA	TTGAGTTGGA	540
CTTAGATCAT	TCCTGTGGGC	ATAGATTTTT	AGGCCGATCC	CTTAAACAGA	AACTGCAAGA	600
TGCGGTGGGG	CAGTGTTTTT	CAATAAAGAA	TTGTAGTGGC	CGACACTCTC	CAGGGCTTCC	660
ATCTAAAAGA	AAGATTCATA	TCAGTGAAGT	CATGTTAGAT	AAGTGCCCTT	TCCCACCTCG	720
CTCAGATTTA	GCCTTTAGGT	GGCATTTTAT	TAAACGACAC	ACTGTTCTTA	TGAGTCCCAA	780
CTCAGATGAA	TGGGTGAGTG	CAGACCTGTC	TGAGAGGAAA	CTGAGAGATG	CTCAGCTGAA	840
ACGAAGAAAC	ACAGAAGATG	ACATAACCCTG	TTTCTCACAT	ACCAATGGCC	AGCCTTGTGT	900
CATAACTGCC	AACAGTGCTT	CGTGACAGG	TGGTCACATA	ACTGGTTCTA	TGATGAACTT	960
GGTCACAAAC	AACAGCATAG	AAGACAGTGA	CATGGATTCA	GAGGATGAAA	TTATAACGCT	1020
GTGCACAAGC	TCCAGAAAAA	GGAATAAGCC	CAGGTGGGAA	ATGGAAGAGG	AGATCCTGCA	1080
GTTGGAGGCA	CCTCCTAAGT	TCCACACCCA	GATCGACTAC	GTCCACTGCC	TTGTTCCAGA	1140
CCTCCTTCAG	ATCAGTAACA	ATCCGTGCTA	CTGGGGTGTC	ATGGACAAAT	ATGCAGCCGA	1200
AGCTCTGCTG	GAAGGAAAGC	CAGAGGGCAC	CTTTTACTT	CGAGATTCAG	CGCAGGAAGA	1260
TTATTTATTC	TCTGTAGT	TTAGACGCTA	CAGTCGTTCT	CTTCATGCTA	GAATTGAGCA	1320
GTGGAATCAT	AAC TTAGCT	TTGATGCCCA	TGATCCTTGT	GTCTTCCATT	CTCCTGATAT	1380

TACTGGGCTC CTGGAACACT ATAAGGACCC CAGTGCCTGT ATGTTCTTTG AGCCGCTCTT 1440  
 GTCCACTCCC TTAATCCGGA CGTTCCCCTT TTCCTTGCGA CATATTTGCA GAACGGTTAT 1500  
 TTGTAATTGT ACGACTTACG ATGGCATCGA TGCCCTTCCC ATTCTTTCGC CTATGAAATT 1560  
 GTATCTGAAG GAATACCATT ATAAATCAAA AGTTAGGTTA CTCAGGATTG ATGTGCCAGA 1620  
 GCAGCAGTGA TGC GGAGAGG TTAGAATGTC GACCTGCATA CATATTTTCA TTTAATATTT 1680  
 TATTTTTCTT ATGCCTCTTT GAATTTTTGT ACAAAGGCAG TTGAATCAAA TAAAACTGTG 1740  
 CCCTAAGTTT TAATTCCAGA TCAATTTATT TTTTTATGA TACACTTGTT ATATATTTTT 1800  
 AAGCAGGTGT TTGGTTTTGT TTTTACCATA TAAATTTACA TATGGTCCAG GCATATTTAC 1860  
 AATTTCAAGG CATTCATAT ACATTTGAAT ATTCTGTATT TTTTAAATAA TCTTTTGTTT 1920  
 TTTCTATGT GTGAAATATT TTGCTAATCT ATGCTATCAG TATTCTTGTA TGACCGAATA 1980  
 GTTACCTATT CTCTTTTCAT CTTGAAGATT TTCAGTAAAG AGTGTGTAA TCAATCCATT 2040  
 ATAATGTAAT TGACTTTTGT AATTTGCCAA TAGGAGTGTT AAACAACAAA ATGATTTAAA 2100  
 ATGAAACTTA ATGTATTTTC ATTTTAAATA TTAATAAAC CAAGTTTGTT TGTTAGTTAT 2160  
 TCTAGCCAAT AAGAAAAGAG AATGTAGCAT CCTAGAGGTG TATTTGTTCT GCAGTTTGGC 2220  
 AGGACCGTCA GTTAGTCCAA ATAAACATCC CCTCAGCGTG GAGGCGAATG GAACCTGTGC 2280  
 TCCTTTCTTA CGGGAAGCTT TGCAAAGCAA AATAGCAGGG TTACAAGCTT GGAGTTGTTA 2340  
 AGGCAACTAG AGTTTTCTCT ATTAATTTAT AGACTGTTGT TGCACCTACT TAGCTCTTTT 2400  
 TTGGGAACTC TAGTCCCAG GGGAAAATAC CTCGTGCC 2438

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ser Gly Gly Gly Pro Trp Arg Ala Gly Gly Gly Ser Gly Lys Ser Asp  
 1 5 10 15



Ser Gly Leu Thr Val Glu Pro Gly Arg Gly Leu Thr Ala Arg Pro Pro  
20 25 30

Pro Gly Gly Ser Arg Thr Arg Ser Gly Ser Gly Arg Ala Ser Leu Pro  
35 40 45

Arg Leu Ser Glu Arg Arg Val Met Ala Val Val Met Ala Ala Gly Ala  
50 55 60

Arg Thr Ala Pro Leu Glu Leu Ser Ser Glu Arg Ser Val Gln Lys Val  
65 70 75 80

Pro Arg Arg Asn Phe Leu Leu Glu Lys Leu Lys Asn Thr Xaa Phe Ile  
85 90 95

Thr Leu Glu Ile Val Lys Asn Leu Phe Lys Met Ala Glu Asn Asn Ser  
100 105 110

Lys Asn Val Asp Val Arg Pro Lys Thr Ser Arg Ser Arg Ser Ala Asp  
115 120 125

Arg Lys Asp Gly Tyr Val Trp Ser Gly Lys Lys Leu Ser Trp Ser Lys  
130 135 140

Lys Ser Glu Ser Cys Ser Glu Ser Glu Ala Ile Gly Thr Val Glu Asn  
145 150 155 160

Val Glu Ile Pro Leu Arg Ser Gln Glu Arg Gln Leu Ser Cys Ser Ser  
165 170 175

Ile Glu Leu Asp Leu Asp His Ser Cys Gly His Arg Phe Leu Gly Arg  
180 185 190

Ser Leu Lys Gln Lys Leu Gln Asp Ala Val Gly Gln Cys Phe Pro Ile  
195 200 205

Lys Asn Cys Ser Gly Arg His Ser Pro Gly Leu Pro Ser Lys Arg Lys  
210 215 220

Ile His Ile Ser Glu Leu Met Leu Asp Lys Cys Pro Phe Pro Pro Arg  
225 230 235 240

Ser Asp Leu Ala Phe Arg Trp His Phe Ile Lys Arg His Thr Val Pro  
245 250 255

Met Ser Pro Asn Ser Asp Glu Trp Val Ser Ala Asp Leu Ser Glu Arg  
260 265 270

Lys Leu Arg Asp Ala Gln Leu Lys Arg Arg Asn Thr Glu Asp Asp Ile  
275 280 285

Pro	Cys	Phe	Ser	His	Thr	Asn	Gly	Gln	Pro	Cys	Val	Ile	Thr	Ala	Asn
290						295					300				
Ser	Ala	Ser	Cys	Thr	Gly	Gly	His	Ile	Thr	Gly	Ser	Met	Met	Asn	Leu
305					310					315					320
Val	Thr	Asn	Asn	Ser	Ile	Glu	Asp	Ser	Asp	Met	Asp	Ser	Glu	Asp	Glu
				325					330					335	
Ile	Ile	Thr	Leu	Cys	Thr	Ser	Ser	Arg	Lys	Arg	Asn	Lys	Pro	Arg	Trp
			340					345					350		
Glu	Met	Glu	Glu	Glu	Ile	Leu	Gln	Leu	Glu	Ala	Pro	Pro	Lys	Phe	His
		355					360					365			
Thr	Gln	Ile	Asp	Tyr	Val	His	Cys	Leu	Val	Pro	Asp	Leu	Leu	Gln	Ile
	370					375					380				
Ser	Asn	Asn	Pro	Cys	Tyr	Trp	Gly	Val	Met	Asp	Lys	Tyr	Ala	Ala	Glu
385					390					395					400
Ala	Leu	Leu	Glu	Gly	Lys	Pro	Glu	Gly	Thr	Phe	Leu	Leu	Arg	Asp	Ser
				405					410					415	
Ala	Gln	Glu	Asp	Tyr	Leu	Phe	Ser	Val	Ser	Phe	Arg	Arg	Tyr	Ser	Arg
			420					425					430		
Ser	Leu	His	Ala	Arg	Ile	Glu	Gln	Trp	Asn	His	Asn	Phe	Ser	Phe	Asp
		435					440					445			
Ala	His	Asp	Pro	Cys	Val	Phe	His	Ser	Pro	Asp	Ile	Thr	Gly	Leu	Leu
	450					455					460				
Glu	His	Tyr	Lys	Asp	Pro	Ser	Ala	Cys	Met	Phe	Phe	Glu	Pro	Leu	Leu
465					470					475					480
Ser	Thr	Pro	Leu	Ile	Arg	Thr	Phe	Pro	Phe	Ser	Leu	Gln	His	Ile	Cys
				485					490					495	
Arg	Thr	Val	Ile	Cys	Asn	Cys	Thr	Thr	Tyr	Asp	Gly	Ile	Asp	Ala	Leu
			500					505					510		
Pro	Ile	Pro	Ser	Pro	Met	Lys	Leu	Tyr	Leu	Lys	Glu	Tyr	His	Tyr	Lys
		515					520					525			
Ser	Lys	Val	Arg	Leu	Leu	Arg	Ile	Asp	Val	Pro	Glu	Gln	Gln		
	530					535					540				

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

all  
cont.

- 165 -

(A) LENGTH: 5000 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CCCTCTGGGC AAGCCGCCCC CCCCCACCC ATCTACCACA CACACACACA CACACACACA	60
CACACATTCA GACCTTGGGG CAAAACAAA GCAAATAAC AACAACAAA AACTGCCTG	120
TGGAAAGTCC TTA CTCAGG AAGGTTGGCA GATGAGGAGC AAGGGAACAT TTTATCAGGA	180
CTGCCACAAA GGAGTCTTTT TTTTAAATGG TTTTCAAGA CAGGGTTTCT CTGTATAGCC	240
CTGGCTGTCC TGGAGCTCAC TTTGTAGACC AGCTGGCCT CGAACTCAGA AATTCGCCTG	300
CCTCTGCCTC CTGAGTGCTG GGATTAAAGG CGTGCAGCAC CATGTCCAAC TGGCATT TTC	360
TCAATTAAGG TTCGTTCTT TCAGATAACT CTAGGTTCTG GGTCAAGCTG ACACAAGGCT	420
ACACAGCACA GTTTGTATGC CACATTCAGT TCAGAAGACA CCAACCTCC CTGGAAGTGG	480
AACTTATGCA CATTTGTGAG CTTCCACTTG GGAGTGGGAA CCTGAAGTGG GTCCTCTGCA	540
AGAGCAGCCG TGCTCTTAAC TGCTGAGCCA TTTCAGCAGC CTCACATCAG AATTAAGTTA	600
GAAATTAGCC GGGTATGAAT CATACCCTTA GAATCCTAGC ATCTGAAAGC AGAGCTAAGA	660
GAAACAGGGA TTCAAGACCA GCTCTTGGCT ACAGAGCCCG TCCTGTCCTA GGATGGGCTA	720
CAAGAGACTA TTTCAAAGCC ATCCAAACAA CAATAACTAC AACAACAACA AGGTTAAAAT	780
TAGGCTGGGC ACAGGGTACA CACCTTTAAT GCCAACACTC AGGAGGCAGA GGCAGGCTGA	840
TCAGTGTGAG TTTGAGTTCA ACGTGGTCTA CATAGGGAGT TCTAGGCCAG CAGAGGT TAC	900
AGTCTCTCTC TCTCTCTCTC TCTCTCTCTC TCTCTCACAC ACACACACAC ACACACACAC	960
ACACACACAC ACACACACGG TGGCATTATG GGATTTTTTT GGGATAAGGT TTCTCTGTCT	1020
AGCCCTGGCA TAGATTC ACT CTGTAGACTA GGCTAGCCTT GAACTCAGAG ATCCGCCTGC	1080
CTCTGCCTCC CAAGTGCTGG GATTATAGGT GTTGCACCAC CACTGCCCAG CCACTTTGGG	1140
ATTTTGAAC TGTATCAAG AGGCTTTCGA GGAGGTCAAA CTCAACAGC AACCTCTCCA	1200
TGATAATGTA GCTAATGATC AAACGACACT CAAA ACTTAA CCCTTAAAGC ACACATCCAC	1260
CAGACAGCGT GCCCACTCGT AGTTCCATTA CTCAGGAGGC TGAAGCAGGA GGATGAAGGA	1320

CTAAGGCTTC	AGCAACCTAG	GGAGCCGCAG	GGGACAGTAG	TCTCAATCCC	TACATTCTCC	1380
TGAACACAGG	AGCAGGAGTT	CAGGAAGGGT	GTCAAGGCCG	CTTACTGATC	TTAGGGCCTC	1440
AGGAATGACT	AGCTCAGGCA	GAGAGAGCAA	AGGTCTCCAG	TGGAGAAGTC	TACACACACA	1500
CACACACACA	CACACACACA	CACACACACA	CAGAATCCAA	GCGGATGACG	TCATCAAAGG	1560
GTTAATTCTA	GTCTGGGATG	GGGGGGAGGG	TGGGGCACGC	AGCTGTCAGG	TGGCTTTTGA	1620
AAAATAAACT	GCTGAAGAGT	CTGACGCCAG	GGAGTCCTGG	GAGGGACAAG	AGGTTACCCA	1680
CTCAAAGAGT	GTGCTCCACA	AAGCATGCGC	GCTTGTCAC	GTCTGGAGTC	GTCACTTATT	1740
TTTTGCCTGG	ATTCTTTGTA	GCCGGTGGGT	TCTCAAGGCG	GTAAGTGGTG	TGGCCGCCGT	1800
GGTCTGGGAG	GTGACGATAG	GGTTAATCGT	CCACAGAGCC	CAGGGGCGGA	GCGCGGGCGG	1860
GCGTCCGCAG	CCCCGCTGGA	GCCGGAAGCA	GTGGCTGGTC	AGGGGCGCTT	CTAGCCTTCC	1920
CTATCTGTAC	TTCCACAGAG	GTCTCTGCGA	GCTAGGGGGA	CAGTGAGGTG	CGGGGTAGGG	1980
GCCCGGCGTT	AGAGCCAGCA	AGGGGACGGT	TCACGGTAAG	GTCTGAGGGA	GAGAGAGCTC	2040
CTGAGAAACT	TGGGGGGCGC	GACACAGATA	GGGTGAAAGC	AGAGTGATAG	ACCTGGGATG	2100
GTTAGGGGAC	CAAGGGAAGA	CCAGGCTGGT	TGGCATAAC	CGGTGAACGG	ATGGGAGTCC	2160
TAGGGAAAGA	TGATGCGCCT	AACAGTCCTT	TCTGTCTCCA	CACCACTCCA	GGGGACGATC	2220
CGGAGCTCAA	CTTTCAAAAG	CGAGACGCCC	CAGCAAGCCT	GTTTTGAGAA	GTTCTTCAGC	2280
GGCTCTCCTC	ATGGGCCAGA	CGGCCCTGGC	AAGGGGCAGC	AGCAGCACCC	CTACCTCGCA	2340
GGCTCTGTAC	TCGGACTTCT	CTCCTCCCGA	GGGCTTGGAG	GAGCTCCTGT	CTGCTCCCCC	2400
TCCTGACCTG	GTTGCCCAAC	GGCACCACGG	CTGGAACCCC	AAGGATTGCT	CCGAGAACAT	2460
CGATGTCAAG	GAAGGGGGTC	TGTGCTTTGA	GCGGCGCCCT	GTGGCCCAGA	GCACTGATGG	2520
AGTCCGGGGG	AAACGGGGCT	ATTCGAGAGG	TCTGCACGCC	TGGGAGATCA	GCTGGCCCCT	2580
GGAGCAAAGG	GGCAACACAG	CCGTGGTGGG	CGTGGCCACC	GCCCTCGCCC	CGCTGCAGGC	2640
TGACCACTAT	GCGGCGCTTT	TGGGCAGCAA	CAGCGAGTCC	TGGGGCTGGG	ATATTGGGCG	2700
GGGAAAATTG	TATCATCAGA	GTAAGGGCCT	CGAGGCCCCC	CAGTATCCAG	CTGGACCTCA	2760
GGGTGAGCAG	CTAGTGGTGC	CAGAGAGACT	GCTGGTGGTT	CTGGACATGG	AGGAGGGGAC	2820
TCTTGGCTAC	TCTATTGGGG	GCACGTACCT	GGGACCAGCC	TTCCGTGGAC	TGAAGGGGAG	2880

GACCCTCTAT	CCCTCTGTAA	GTGCTGTTTG	GGGCCAGTGC	CAGGTCCGCA	TCCGCTACAT	2940
GGGCGAAAGA	AGAGGTGAGA	TACGGACTAG	GTGTGGGGAG	ATCACTACTC	TTGGCAATGG	3000
TTTGGGCTGG	AAACTCATGG	TTGGAGCACA	GGAAGTAGGC	TTCTTGTCAC	TTTGGCCTGT	3060
CACTTAGATG	GCCTTGATC	TAGCTTCACT	CCCAATCCCT	ATTGGATGTG	ATGCACAAAT	3120
TCAGAGCCTT	TGGGTCTCCC	TCAGCTGAGG	TGGCGGTGGA	AATGGAGGAA	GAAGGAAGGG	3180
TGCCTGAGCA	GGATCTCAAG	TTCAAGGATG	CCTGGAGTTG	CTTACTTACC	TTGTCTTCCT	3240
TCTCTCTCCG	CAGTGGAGGA	ACCACAATCC	CTTCTGCACC	TGAGCCGCCT	GTGTGTGCGC	3300
CATGCTCTGG	GGGACACCCG	GCTGGGTCAA	ATATCCACTC	TGCCTTTGCC	CCCTGCCATG	3360
AAGCGCTATC	TGCTCTACAA	ATGACCCAGT	AGTACAGGGT	GTGCTGGCAC	CCTACCGTGG	3420
GGACAGGTGG	AGAGGCACCC	GCTGGCCTAG	ACAACCTTTAA	AAAGCTGGTG	AAGCTGGGGG	3480
GGGGGGGCTG	GACCCCTTCA	CCTCCCCTTC	TCACAGGAGC	AAGACATATA	GAAATGATAT	3540
TAAACACCAT	GGCAGCCTGG	GACAAAGAGG	TTTTTTGAAGT	AAAAAATGAG	ATGTATTGTC	3600
ACAACCTGTT	TCATTATTGT	TTTTTTGTTTT	GTTTTTACACT	CCCCCACCCC	AGGCTAGAGC	3660
CCCATCACTG	TCTTAAGGAA	TTATGACAAC	CCACAAAGCT	CAGGCCCAGG	TGTTTATTTT	3720
CCTTACATGT	AGGATGGTTC	ACAAACACAA	TACAGGGGCT	TTGGCACCGT	GGGGGAGGGG	3780
ACTATCCCAG	GCCTCTTAGG	GTCTCATGTA	TACCGAATTC	AGACCCGAAA	GCTCTGAATT	3840
TCTGCATCAG	ACATCCAGTA	GAACTTGGGA	GTGAAGCTAG	AGCCAAGGCC	ATCTAAGTGA	3900
CAGGCCAAAG	TGACACGAAG	CCCACTTCCT	GTGCTCCAAC	CATGAGTTTC	CAGCCCAAAC	3960
CAATGGAAGG	TGATTTCACT	TGTCAGGGCC	CAAAGGGACA	GTCAGTTCTA	CTCCCTCCCC	4020
TCTACTAGGAG	CCACCTTGGT	GACAGTTGAT	TCTACCCACT	GTAAGTGGTA	AAGGGATTGG	4080
CCTGGTCCCA	ACCATAATAG	GGCGGTGGAA	ACGGCTCAGG	AGGGTACAGC	GTGGATTAGG	4140
CCACAAGATG	GGGCAGATGA	TGTCATCAGA	AGCATGTGAC	CGGTGGGAGC	AGTTACTAAA	4200
CTTCTGGGCA	ACCTAGTCCA	TGCTATGCAG	GCAGGTAGAG	GGATGGGCAG	TGCTCATTTG	4260
TTGGCATTGA	TGATGTCCAC	AAATTCAGGC	TTGAGAGATG	CGCCACCCAC	AAGGAAGCCG	4320
TCCACGTCAG	GCTGGCTTGC	CAGCTCTTTG	CAGGTTGCTC	CAGTCACAGA	ACCTGTACCA	4380
GGAACAAGAA	GACAGTTTGG	TCAGGTCTAT	GATCAGAACA	CTTAAGCCCC	ACCTCTCTGT	4440

GCAAGGCAGC CTCAGTCTGT CTTAGCCCAT TTCCGTCTTA GCTAGAGCCA AAGCCACTCA 4500  
 CCTCCATAAA TGATCCGGGT GCTCTGAGCC ACCCCATCAT TGACATTGGA TTTCAGCCAT 4560  
 CCCC GGAGCT TCTCGTGAC TTCCTGTGCC TAGAAGGAGG AEGCAGAGCT ACTAAGTAAG 4620  
 CTCCTTCCTA TCTATCATTC AAGGAGTAAA AACCCTGGT TCTCACATAG AGTTGAGTTT 4680  
 CCAGAAAAGC CCCGGGACCA GAGAGTGGCA AGGCTCCAAT CCCACCAGGC TTGGAATGAA 4740  
 CATTTTTGGC AAAGTCACTC TCCTTGGTGA GTTTGGGGGC CCTCTGTCTC TAAAGGGGCT 4800  
 TGGATGGGCT CCATAGCTGT GTGAGTCTGT TAAAGCCGGA CAGGCTGAGG AGCTCTGGGT 4860  
 AGTTACCTGC TGAGGGGTTG CCGTCTTGCC AGTCCAATG GCCCACACAG GTTCATAGGC 4920  
 CAGGACCACC TTGCTCCAGT CTTTCACATT ATCTGTGGGG CAGAGAGGAG AGTGAGTAGG 4980  
 AAGGAGCTGA CCCGCCAAGC 5000

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Met	Gly	Gln	Thr	Ala	Leu	Ala	Arg	Gly	Ser	Ser	Ser	Thr	Pro	Thr	Ser	1	5	10	15
Gln	Ala	Leu	Tyr	Ser	Asp	Phe	Ser	Pro	Pro	Glu	Gly	Leu	Glu	Glu	Leu	20	25	30	
Leu	Ser	Ala	Pro	Pro	Pro	Asp	Leu	Val	Ala	Gln	Arg	His	His	Gly	Trp	35	40	45	
Asn	Pro	Lys	Asp	Cys	Ser	Glu	Asn	Ile	Asp	Val	Lys	Glu	Gly	Gly	Leu	50	55	60	
Cys	Phe	Glu	Arg	Arg	Pro	Val	Ala	Gln	Ser	Thr	Asp	Gly	Val	Arg	Gly	65	70	75	80
Lys	Arg	Gly	Tyr	Ser	Arg	Gly	Leu	His	Ala	Trp	Glu	Ile	Ser	Trp	Pro	85	90	95	
Leu	Glu	Gln	Arg	Gly	Thr	His	Ala	Val	Val	Gly	Val	Ala	Thr	Ala	Leu	100	105	110	

Ala Pro Leu Gln Ala Asp His Tyr Ala Ala Leu Leu Gly Ser Asn Ser  
115 120 125

Glu Ser Trp Gly Trp Asp Ile Gly Arg Gly Lys Leu Tyr His Gln Ser  
130 135 140

Lys Gly Leu Glu Ala Pro Gln Tyr Pro Ala Gly Pro Gln Gly Glu Gln  
145 150 155 160

Leu Val Val Pro Glu Arg Leu Leu Val Val Leu Asp Met Glu Glu Gly  
165 170 175

Thr Leu Gly Tyr Ser Ile Gly Gly Thr Tyr Leu Gly Pro Ala Phe Arg  
180 185 190

Gly Leu Lys Gly Arg Thr Leu Tyr Pro Ser Val Ser Ala Val Trp Gly  
195 200 205

Gln Cys Gln Val Arg Ile Arg Tyr Met Gly Glu Arg Arg Val Glu Glu  
210 215 220

Pro Gln Ser Leu Leu His Leu Ser Arg Leu Cys Val Arg His Ala Leu  
225 230 235 240

Gly Asp Thr Arg Leu Gly Gln Ile Ser Thr Leu Pro Leu Pro Pro Ala  
245 250 255

Met Lys Arg Tyr Leu Leu Tyr Lys  
260

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5615 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

GTACTTTCTT TATATCTCCA TAATTTTATT TACTATTACT ACATGATACA TTATTTTATA	60
AAAGTCTTTG TAACCTCCTT AAGGATTCAC TGCTTAATCT CCAGTGCTTA GCACAAATCA	120
TTAAATGCGA ACCAGAACT CTCCAAATG TGTTACATCT ATAACCTCAT TGGATTCTCA	180
CTACCAACCC CATGCAATAG ATACTAATGT GATCTCTGTC TTACAGAGGA AGAAACAGGC	240
ACAGGGAGGT TCAGTAATTT GCCCAAGGTC ATACACACAC TGGCCTTCAG GTATTCATGC	300

CCGGGGAGTC	TGGTCCCACA	GCTGGCATGT	TTGCCATTAT	ATTATATTGC	CTCCTTATAG	360
TGTCGGCACT	CATTAAGCAC	ATTGACAGCT	ATGCTTGGTG	AGTGAETACT	ATGTACCCAG	420
CTCTGTGCTA	CATGCTTTAC	CTGGATTATT	TCAACTGCAC	AAQAACCCTG	TGAGGTAAGT	480
ACCATCATTG	CTCCTATTTT	ACATAACAGA	AAACTACAGA	AATCTGGGGC	TGGGCGTAGT	540
GGCTCATGCC	TGAAATCCCA	GCACTTTGGG	AGACCCTGTC	TCTAAAAAAA	ATTTTTTTTT	600
GGCCGGACGT	GGTGGCTCAC	ACCTGTAATC	TCAGCACTTT	GGGAGGCTAA	GGCAGGCAGA	660
TCACAAGGTC	AGGAGTTCTA	GACCAGCCTG	GCCAACATGG	CAAAACCCTG	TGTCTACTAA	720
AAATACAAAA	AATAGCTAGG	CGTGGTGCCA	GGTGCCTGTA	ATCCCAGCTA	CTCAGGAGGC	780
TGAGGCAGGA	GAATCCCCTG	AACCTGGGAG	ATGGAGGTTA	CAGAGAGCCG	AGATCGTGCC	840
GCTGCACTCC	AGCCTGGGCA	ACAAGAGCAA	GACTCTGTCT	CGAAAAAAT	AAAAATAAAA	900
ATAAAAAATAT	TTTTTTAAAA	ATTAGCTGGG	TGTGGTAGCA	CATGCCTGTA	GTCCCAGCTA	960
CTTGGGAGGC	TGAGGTAGGA	GGATCACTTG	AGCCCAGGAG	GTCAAGGCTG	CAGTGGGCTG	1020
TGATGGCGCC	ACTGCACTCT	AGCCTTGGTG	ACAGCAAGAC	CCTGTCTCAA	AAAAAAAAAA	1080
AAGAGAAATC	GGGCAACTTC	CCCAAGATCG	CGCAGTTAAC	TAGTGGCATA	GCTTCACTCA	1140
AACTCGAAGT	CTTAATCAGG	ACACTCTACC	AAATGAGATC	AACGGCTCAG	TAATGGATTG	1200
GCATCCAGTA	TGAAGACTGG	ACCAGCAGGG	AGAACTATGA	TGCGTACAGC	CTAGAGCCTG	1260
AAGCAGATTT	CACAGCCTCA	GAGGTGGCAC	AGGCTGACTC	ACAACCCGGG	GCAGAAAGGG	1320
ACCAGCCCAG	AAACAGTGAC	CCAGAATCAC	AGGGAAGTAG	AAATGGGATT	CGGCACAATG	1380
AAGCCCCCTC	TTGAACCCAT	GCTCCTTACC	CTCAGGGGCG	CAGGAGTTAG	TCGCTCAGGC	1440
GGCTCAAAGG	TCTTGACGGT	GGAGAACACC	ATCCCCAGGG	ATTCCCGACG	CGGTGATGCC	1500
ATCAAAGCGT	TAATTCTGAG	ATGGGCCTGC	CCGGGTGCGG	ACTCTGCCGC	AGCAAGAGAA	1560
GGGTAACTG	CCCCGGGCCT	TCGCCGTGGG	GGCGGGGCCT	CGGGGAGGGT	CACAGCCCGG	1620
GACTGAGACC	CGAGGTTAAC	CGCCCGGGGT	GGGCTCCACG	GGGGCGGGGC	ATGCTCTCCG	1680
CGGCTGCTGC	CGGTATAGAG	CGGTAAGTGC	CCAGGAGGGG	GCGGGGCCCC	ACAGGGGCGT	1740
GGCCTCGGAG	CTGCACGGCC	GTGGGCGGCG	ATGAGAGGGT	TAAGCCCCAG	AGGGCCCTGG	1800
AGGGGCGGGG	CCGCGGGACG	GGCTCGGCCC	AAGGGAGGAG	CTGGGGGCGG	AAGCGGCCGG	1860



OH  
CONF

- 171 -

CGGTCTGCGC CCTGCGCGCC TCGGCTTCTT TCCGCCC GGC TCCTTCAGAG GCCCGGCGAC 1920  
CTCCAGGGCT GGAAGTCAA CCGAGGTTCG GGGGAGCGG CGAGGGCTCC GGGCGAGTAA 1980  
GGGGGATGGT CCATGCTGAG GCCCAAATGG GGC GAAC TCG CGAGAGTCTC TGGCGACCTG 2040  
GATCAGATGG GGCGAGGGCA GATGAAGGGC CCAGGAGCTT TGGGGCAGCG AGGAGGGAGG 2100  
AGCGGGCCCCG TTGGCAAAC TGGGTGAAAG GATGGGGTAC CTGGGTGACG AGCCCCCGCC 2160  
AGGATTCTGC TCTTCACGCC CCTTTTCTCC CAGCTCCCTT CCAGGTCAAT CCAAAC TGG 2220  
GCTCAACTTT CAGAAGAGAA AGACGCCCCA GCAAGCCTCT TTCGGGGAGT CCTCTAGCTC 2280  
CTCACCTCCA TGGGCCAGAC AGCTCTGGCA GGGGGCAGCA GCAGACCCCC CACGCCACAG 2340  
GCCCTGTACC CTGACCTCTC CTGTCCCGAG GGCTTGAAG AGCTGCTGTC TGCACCCCCT 2400  
CCTGACCTGG GGGCCAGCG GCGCCACGGT TGAACCCCA AAGACTGTTC AGAGAACATC 2460  
GAGGTCAAGG AAGGAGGGTT GTACTTTGAG CGGCGGCCCG TGGCCCAGAG CACTGATGGG 2520  
GCCCCGGGGTA AGAGGGGCTA TTCAAGGGGC CTGCACGCCT GGGAGATCAG CTGGCCCCCTA 2580  
GAGCAGAGGG GCACGCATGC CGTGGTGGGC GTGGCCACGG CCCTCGCCCC GCTGCAGACT 2640  
GACCACTACG CGGCGCTGCT GGGCAGCAAC AGCGAGTCGT GGGGCTGGGA CATCGGGCGG 2700  
GGGAAGCTGT ACCATCAGAG CAAGGGGGCC GGAGCCCCC AGTATCCAGC GGGAAC TCA 2760  
GGTGAGCAGC TGGAGGTGCC AGAGAGACTG CTGGTGGTTC TGGACATGGA GGAGGGAAC T 2820  
CTGGGCTACG CTATTGGGGG CACCTACCTG GGGCCAGCAT TCCGCGGACT GAAGGGCAGG 2880  
ACCCTCTATC CGGCAGTAAG CGCTGTCTGG GGCCAGTGCC AGGTCCGCAT CCGCTACCTG 2940  
GGCGAAAGGA GAGGTGAGGC CTGGGGCAGA CGTGGGGAGA ACTTTCTGTC CCTGGTGGCA 3000  
GTGGTTTGGG ATGGAAACTC TTCTGACAAG AGCAGAGGGG ATGGACCTTC ATCCAGCCTG 3060  
CCTCAACCTC TGTTCACTGC TGGGAAAGGC TAGGGGTCTT CACAGCTGTT ATTTAATTTA 3120  
ACCCAACAGC AATAGAGGTG AAACAGGCTT GAGAAAGCAA CTTTCTCAAG TTCTCTTGGC 3180  
CAGTAAATGG TGAACCTTCA GAATGGAGGG AGGAACTGCA GGGATGAGAG AATTCAGGAG 3240  
ATATCAACCC CTGAGCAAGA GGTGCAAAGC GTTAGGTACT GGGTTTGATG TACAGGTCCA 3300  
AAAGAAGGAT GGGCAGAGCC AGGTACCCAG GCTGTATACC GGATTCCCTG GGCTCTAACC 3360  
TGTCTCTGTG CCACATACCT ACTTCCTTCC TCAGCCACAC CTCTGGATGG AGACACTGGG 3420

GCCCTGGGCA	CCAGGGAGGA	GAGCAGTGG	GGAGGCAGG	CCTTAGGGTG	GGGCAGCAGG	3480
GGAGGAGCCT	CCCCAGGAAC	TGACTGGGTC	CAGGGCTTGG	AGCTGCTCTC	TGCAGTTGTG	3540
TGGGCTGTAG	AGTGGAGGGC	CATCCCTCCT	CACCTCAGCC	CGAGCTCCCA	AGCCTCTGGA	3600
GTCAAAGCCT	GGGCCAGCTC	CACCACTGTC	AGAGCCACCT	TGGCCTGTTG	TTTAGAGGGC	3660
CTTAGCCAGC	TCTTCACCCC	CAGCTCTGAC	TAGGGATGTG	TGAAATCTTA	TCTGGGAGGC	3720
AGAACTTCCG	GGTATCTCAA	ATTCCCCTTT	CAGCCAGGTG	GGCACACTCG	AAGCAGGAAA	3780
GCAGAAAGGC	ATCTGAGTAG	GACCCCGTAG	TTTGAGGACA	TCTGGCTGGT	GGCTGCACCC	3840
ATACTTAÇAT	TCCCCTCCTT	CTCTCTCCCA	GCGGAGCCAC	ACTCCCTTCT	GCACCTGAGC	3900
CGCCTGTGTG	TGCGCCACAA	CCTGGGGGAT	AQCCGGCTCG	GCCAGGTGTC	TGCCCCTGCCC	3960
TTGCCCCCTG	CCATGAAGCG	CTACCTGCTC	TACCAGTGAG	CCCTGTGATA	CCACAGACTG	4020
TGCTGAGGTC	TTGCCACCAC	CCCTCCCCTT	GGGGAGGTGG	GGAGGCACTG	CTGGCCTAGA	4080
CCAGCTGCTG	AAAGCTGGTG	AGGCTGAGCC	CCTACCCCAA	CCCAAGCTCT	GCGGAAATCA	4140
ACAGCCCCAG	AGCCACTTGG	AGGGAGGAAG	AAAGGGAGCC	GGCGTTCAAG	GCTATGACAG	4200
TCTGCTACGC	AAAACATTTT	TTCAAGTAAA	AATAGTAAGA	GATGTTGTTA	TAGAAACCTG	4260
TTCTTGTTTT	TTTTTTTTTC	TTGCACAAAT	GATCATTTAT	ATAGCTGCCT	CAAAAAGGAA	4320
GATTATCTGG	GCAAGTCCAG	TGAAGGCAGA	CAAACCACAA	GACCTAGTGC	CAGGTTTATT	4380
CCCTCACATG	GGTGGTTCAC	ATACACAGCA	CAGAGGCACG	GGCACCATGG	GAGAGGGCAG	4440
CACTCCTGCC	TTCTGAGGGG	ATCTTGGCCT	CACGGTGTA	GAAGGGAGAG	GATGGTTTCT	4500
CTTCTGCCCT	CACTAGGGCC	TAGGGAACCC	AGGAGCAAAT	CCCACCACGC	CTTCCATCTC	4560
TCAGCCAAGG	AGAAGCCACC	TTGGTGACGT	TTAGTTCCAA	CCATTATAGT	AAGTGGAGAA	4620
GGGATTGGCC	TGGTCCCAAC	CATTACAGGG	TGAAGATATA	AACAGTAAAG	GAAGATACAG	4680
TTTGGATGAG	GCCACAGGAA	GGAGCAGATG	ACACCATCAG	AAGCATATGC	AGGGAAAGGG	4740
CAGTTACTGG	GCTTCTGGGC	TGCTTAGTCC	CTGGCTTGGC	AGGAAGGGTA	GGGAAGATGG	4800
ATGGGGCTCA	TTGTTTGGCA	TTGATGATGT	CCACGAATTC	GGGCTTGAGG	GAAGCACCAC	4860
CCACAAGGAA	GCCATCCACA	TCAGGCTGGC	TGGCCAGCTC	CTTGCAGGTT	GCCCCAGTCA	4920
CAGAGCCTGG	GAAGGGAGCA	GAACAAGGGC	TTGGTCAAGA	ATGGGATGAG	TCTGCCCCAT	4980

B1

- 173 -

CCCCACCTCC ATGTCCGAGG GCTCAGTCTA GTCCTCAGCC CACTCCACCT CAGCCGGGAA 5040  
 CCAAAGCCAC TCACCTCCAT AAATGATACG GGTGCTCTGA GCCACCGCAT CAGAGACGTT 5100  
 GGACTTCAGC CATCCTCGGA GCTTCTCGTG TACTTCCTGG GCCTAGAACA AGAAGCTGGC 5160  
 CTAAGTAAGA CCTTTTCTGC CTCTCTAAGA GGAAAAATCA CTGGCACCAG TGGACACTTA 5220  
 GTGTGGTTTC TGAAGTGGTC AGAGTACCAG GGCTCTGATC CAAGCCAGGC CCTGGACTGG 5280  
 ATGCCCTTGG ACAAGTCACT GTCTCTGGGT TCAAGGTCTC TGTGTCTTTG AAATAAGGGG 5340  
 TTGCCCCATG TGGGCTGTGT CTGTCCAAAC CTATTGAGGC AGGCTGGGAT GAGGGCAGGG 5400  
 CTCCTGGGCC CGGTTACCTG TTGGGGTGTG GCAGTCTTGC CAGTACCAAT GGCCCACACA 5460  
 GGCTCATAGG CCAGGACGAC CTTGCTCCAG TCCTTCACGT TATCTGCAGG GCAGAGATAC 5520  
 AGATGGAGGG AAGGGTGAAC AAGAAAGAGC TCTCCAGCCA GGTTCCTCCG AGTACGAAGA 5580  
 ACGGTGGCCT ACTGCCCCCT AGTGGACATT GGGGG 5615

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met	Gly	Gln	Thr	Ala	Leu	Ala	Gly	Gly	Ser	Ser	Ser	Thr	Pro	Thr	Pro
1				5					10					15	
Gln	Ala	Leu	Tyr	Pro	Asp	Leu	Ser	Cys	Pro	Glu	Gly	Leu	Glu	Glu	Leu
			20					25					30		
Leu	Ser	Ala	Pro	Pro	Pro	Asp	Leu	Gly	Ala	Gln	Arg	Arg	His	Gly	Trp
		35					40					45			
Asn	Pro	Lys	Asp	Cys	Ser	Glu	Asn	Ile	Glu	Val	Lys	Glu	Gly	Gly	Leu
		50				55					60				
Tyr	Phe	Glu	Arg	Arg	Pro	Val	Ala	Gln	Ser	Thr	Asp	Gly	Ala	Arg	Gly
65					70					75				80	
Lys	Arg	Gly	Tyr	Ser	Arg	Gly	Leu	His	Ala	Trp	Glu	Ile	Ser	Trp	Pro
				85					90					95	

B1 cont

- 174 -

Leu	Glu	Gln	Arg	Gly	Thr	His	Ala	Val	Val	Gly	Val	Ala	Thr	Ala	Leu
			100					105					110		
Ala	Pro	Leu	Gln	Thr	Asp	His	Tyr	Ala	Ala	Leu	Leu	Gly	Ser	Asn	Ser
		115					120					125			
Glu	Ser	Trp	Gly	Trp	Asp	Ile	Gly	Arg	Gly	Lys	Leu	Tyr	His	Gln	Ser
		130				135					140				
Lys	Gly	Pro	Gly	Ala	Pro	Gln	Tyr	Pro	Ala	Gly	Thr	Gln	Gly	Glu	Gln
145					150					155					160
Leu	Glu	Val	Pro	Glu	Arg	Leu	Leu	Val	Val	Leu	Asp	Met	Glu	Glu	Gly
			165					170						175	
Thr	Leu	Gly	Tyr	Ala	Ile	Gly	Gly	Thr	Tyr	Leu	Gly	Pro	Ala	Phe	Arg
		180						185					190		
Gly	Leu	Lys	Gly	Arg	Thr	Leu	Tyr	Pro	Ala	Val	Ser	Ala	Val	Trp	Gly
		195					200					205			
Gln	Cys	Gln	Val	Arg	Ile	Arg	Tyr	Leu	Gly	Glu	Arg	Arg	Ala	Glu	Pro
	210					215					220				
His	Ser	Leu	Leu	His	Leu	Ser	Arg	Leu	Cys	Val	Arg	His	Asn	Leu	Gly
225					230					235					240
Asp	Thr	Arg	Leu	Gly	Gln	Val	Ser	Ala	Leu	Pro	Leu	Pro	Pro	Ala	Met
			245						250					255	
Lys	Arg	Tyr	Leu	Leu	Tyr	Gln									
			260												

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGCTAGATCT GGACCCTACA ATGGCAGC

28

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid

7/15  
conclude

8/1  
conclude

- 175 -

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGCTAGATCT GCCATCCTAC TCGAGGGGCC AGCTGG